

Functional delta-method for the bootstrap of uniformly quasi-Hadamard differentiable functionals

Eric Beutner*

Henryk Zähle†

Abstract

The functional delta-method provides a convenient tool for deriving bootstrap consistency of a sequence of plug-in estimators w.r.t. a given functional from bootstrap consistency of the underlying sequence of estimators. It has recently been shown in [7] that the range of applications of the functional delta-method for establishing bootstrap consistency *in probability* of the sequence of plug-in estimators can be considerably enlarged by replacing the usual condition of Hadamard differentiability of the given functional by the weaker condition of quasi-Hadamard differentiability. Here we introduce the notion of uniform quasi-Hadamard differentiability and show that this notion extends the set of functionals for which *almost sure* bootstrap consistency of the corresponding sequence of plug-in estimators can be obtained by the functional delta-method. We illustrate the benefit of our results by means of the Average Value at Risk functional as well as the composition of the Average Value at Risk functional and the compound convolution functional. For the latter we use a chain rule to be proved here. In our examples we consider the weighted exchangeable bootstrap for independent observations and the blockwise bootstrap for β -mixing observations.

Keywords: Bootstrap; Functional delta-method; Uniform quasi-Hadamard differentiability; Chain rule; Statistical functional; Weak convergence for the open-ball σ -algebra; Average Value at Risk; Compound distribution; Weighted exchangeable bootstrap; Blockwise bootstrap

*Department of Quant. Economics, Maastricht University, e.beutner@maastrichtuniversity.nl

†Department of Mathematics, Saarland University, zaehle@math.uni-sb.de

1. Introduction

The functional delta-method is a widely used technique to derive bootstrap consistency for a sequence of plug-in estimators w.r.t. a map H from bootstrap consistency of the underlying sequence of estimators. An essential limitation of the classical functional delta-method for proving bootstrap consistency in probability (or outer probability) is the condition of Hadamard differentiability on H (cf. Theorem 3.9.11 of [29]). It is commonly acknowledged that Hadamard differentiability fails for many relevant maps H . Recently, it was demonstrated in [7] that a functional delta-method for the bootstrap *in probability* can also be proved for *quasi*-Hadamard differentiable maps H . Quasi-Hadamard differentiability is a weaker notion of “differentiability” than Hadamard differentiability and can be obtained for many relevant statistical functionals H ; see, e.g., [4, 5, 6, 18, 19]. Using the classical functional delta-method to prove almost sure (or outer almost sure) bootstrap consistency for a sequence of plug-in estimators w.r.t. a map H from almost sure (or outer almost sure) bootstrap consistency of the underlying sequence of estimators requires *uniform* Hadamard differentiability on H (cf. Theorem 3.9.11 of [29]). In the present article we will introduce the notion of *uniform quasi*-Hadamard differentiability and demonstrate that one can even obtain a functional delta-method for the *almost sure* bootstrap and *uniformly quasi*-Hadamard differentiable maps H . Proposition 4.1 below shows that the notion of uniform quasi-Hadamard differentiability is weaker than uniform Hadamard differentiability, because this proposition shows that the Average Value at Risk functional, which fails to be Hadamard differentiable, is uniformly quasi-Hadamard differentiable.

To explain the background and the contribution of the paper at hand more precisely, assume that we are given an estimator \hat{T}_n for a parameter θ in a vector space, with n denoting the sample size, and that we are actually interested in the aspect $H(\theta)$ of θ . Here H is any map taking values in a vector space. Then $H(\hat{T}_n)$ is often a reasonable estimator for $H(\theta)$. One of the main objects in statistical inference is the distribution of the error $H(\hat{T}_n) - H(\theta)$, because the error distribution can theoretically be used to derive confidence regions for $H(\theta)$. However in applications the exact specification of the error distribution is often hardly possible or even impossible. A widely used way out is to derive the *asymptotic* error distribution, i.e. the weak limit μ of $\text{law}\{a_n(H(\hat{T}_n) - H(\theta))\}$ for suitable normalizing constants a_n tending to infinity, and to use μ as an approximation for $\mu_n := \text{law}\{a_n(H(\hat{T}_n) - H(\theta))\}$ for large n . Since μ usually still depends on the unknown parameter θ , one should use the notation μ_θ instead of μ . In particular, one actually uses $\mu_{\hat{T}_n} := \mu_\theta|_{\theta=\hat{T}_n}$ as an approximation for μ_n for large n .

Not least because of the estimation of the parameter θ of μ_θ , the approximation of μ_n by $\mu_{\hat{T}_n}$ is typically only moderate. An often more efficient alternative technique to approximate μ_n is the bootstrap. The bootstrap has been introduced by Efron [14] in 1979 and many variants of his method have been introduced since then. One may refer to

[11, 15, 20, 27] for general accounts on this topic. The basic idea of the bootstrap is the following. Re-sampling the original sample according to a certain re-sampling mechanism (depending on the particular bootstrap method) one can sometimes construct a so-called bootstrap version \hat{T}_n^* of \hat{T}_n for which the conditional law of $a_n(H(\hat{T}_n^*) - H(\hat{T}_n))$ “given the sample” has the same weak limit μ_θ as the law of $a_n(H(\hat{T}_n) - H(\theta))$ has. The latter is referred to as bootstrap consistency. Since \hat{T}_n^* depends only on the sample and the re-sampling mechanism, one can at least numerically determine the conditional law of $a_n(H(\hat{T}_n^*) - H(\hat{T}_n))$ “given the sample” by means of a Monte Carlo simulation based on $L \gg n$ repetitions. The resulting law μ_L^* can then be used as an approximation of μ_n , at least for large n .

In applications the roles of θ and \hat{T}_n are often played by a distribution function F and the empirical distribution function \hat{F}_n of n random variables that are identically distributed according to F , respectively. Not least for this particular setting several results on bootstrap consistency for \hat{T}_n are known (see also Section 3). The functional delta-method then ensures that bootstrap consistency also holds for $H(\hat{T}_n)$ when H is suitably differentiable at θ . Technically speaking, as indicated above, one has to distinguish between two types of bootstrap consistency. First bootstrap consistency *in probability* for $H(\hat{T}_n)$ can be associated with

$$\lim_{n \rightarrow \infty} \mathbb{P}^{\text{out}}[\{\omega \in \Omega : d_{\text{BL}}^\circ(P_n(\omega, \cdot), \mu_\theta) \geq \delta\}] = 0 \quad \text{for all } \delta > 0, \quad (1)$$

where ω represents the sample, $P_n(\omega, \cdot)$ denotes the conditional law of $a_n(H(\hat{T}_n^*) - H(\hat{T}_n))$ given the sample ω , d_{BL}° is the bounded Lipschitz distance, and the superscript ^{out} refers to outer probability. At this point it is worth pointing out that we consider weak convergence (resp. convergence in distribution) w.r.t. the open-ball σ -algebra, in symbols \Rightarrow° (resp. \leadsto°), as defined in [8, Section 6] (see also [12, 13, 24, 28]) and that by the Portmanteau theorem A.3 in [7] weak convergence $\mu_n \Rightarrow^\circ \mu$ holds if and only if $d_{\text{BL}}^\circ(\mu_n, \mu) \rightarrow 0$. Second bootstrap consistency *almost surely* for $H(\hat{T}_n)$ means that

$$\text{law}\{a_n(H(\hat{T}_n^*(\omega, \cdot)) - H(\hat{T}_n(\omega)))\} \Rightarrow^\circ \mu_\theta \quad \mathbb{P}\text{-a.e. } \omega. \quad (2)$$

In [7] it has been shown that (1) follows from the respective analogue for \hat{T}_n when H is suitably quasi-Hadamard differentiable at θ . This extends Theorem 3.9.11 of [29] which covers only Hadamard differentiable maps. In this article we will show that (2) follows from the respective analogue for \hat{T}_n when H is suitably *uniformly* quasi-Hadamard differentiable at θ ; the notion of uniform quasi-Hadamard differentiable will be introduced in Definition 2.1 below. This extends Theorem 3.9.13 of [29] which covers only Hadamard differentiable maps.

To demonstrate that the theory presented here leads directly to new results for interesting applications we consider the Average Value at Risk functional and the compound distribution functional. To the best of our knowledge so far there do not exist results on

almost sure bootstrap consistency for the Average Value at Risk functional when the underlying data are dependent. The same seems to be true for the compound distribution functional and consequently also for the composition of the Average Value at Risk functional and the compound distribution functional.

The rest of the article is organized as follows. In Section 2 we introduce the definition of uniform quasi-Hadamard differentiability and prove a functional delta-method for almost sure bootstrap consistency based on it. In Section 3 this functional delta-method is discussed if the underlying sequence of estimators is the empirical distribution function. Section 4 shows that the Average Value at Risk functional and the compound distribution functional are uniformly quasi-Hadamard differentiable. Moreover, we show there using a chain rule that the composition of the Average Value at Risk functional and the compound distribution functional is uniformly quasi-Hadamard differentiable. This chain rule is proved in the Appendix A.2 where we also prove a delta-method for uniformly quasi-Hadamard differentiable maps that is the basis for the main result of Section 2. In the Appendix A.1 we give results on convergence in distribution for the open-ball σ -algebra which are needed for the main results.

2. Abstract delta-method for the bootstrap

Theorem 2.3 below provides an abstract delta-method for the almost sure bootstrap. It is based on the notion of uniform quasi-Hadamard differentiability which we introduce first. This sort of differentiability extends the notion of quasi-Hadamard differentiability as introduced in [5, 7]. The latter corresponds to the differentiability concept in (i) of Definition 2.1 ahead with \mathcal{S} and $\tilde{\mathbf{E}}$ as in (iii) and (v) of this definition. Let \mathbf{V} and $\tilde{\mathbf{V}}$ be vector spaces. Let $\mathbf{E} \subseteq \mathbf{V}$ and $\tilde{\mathbf{E}} \subseteq \tilde{\mathbf{V}}$ be subspaces equipped with norms $\|\cdot\|_{\mathbf{E}}$ and $\|\cdot\|_{\tilde{\mathbf{E}}}$, respectively. Let

$$H : \mathbf{V}_H \longrightarrow \tilde{\mathbf{V}}$$

be any map defined on some subset $\mathbf{V}_H \subseteq \mathbf{V}$.

Definition 2.1 *Let \mathbf{E}_0 be a subset of \mathbf{E} , and \mathcal{S} be a set of sequences in \mathbf{V}_H .*

(i) The map H is said to be uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\tilde{\mathbf{E}}$ if $H(y_1) - H(y_2) \in \tilde{\mathbf{E}}$ for all $y_1, y_2 \in \mathbf{V}_H$, $n \in \mathbb{N}$, and there is some continuous map $\dot{H}_{\mathcal{S}} : \mathbf{E}_0 \rightarrow \tilde{\mathbf{E}}$ such that

$$\lim_{n \rightarrow \infty} \left\| \dot{H}_{\mathcal{S}}(x) - \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n} \right\|_{\tilde{\mathbf{E}}} = 0 \quad (3)$$

holds for each quadruple $((\theta_n), x, (x_n), (\varepsilon_n))$, with $(\theta_n) \in \mathcal{S}$, $x \in \mathbf{E}_0$, $(x_n) \subseteq \mathbf{E}$ satisfying $\|x_n - x\|_{\mathbf{E}} \rightarrow 0$ as well as $(\theta_n + \varepsilon_n x_n) \subseteq \mathbf{V}_H$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \rightarrow 0$. In this case the map $\dot{H}_{\mathcal{S}}$ is called uniform quasi-Hadamard derivative of H w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$.

(ii) If \mathcal{S} consists of all sequences $(\theta_n) \subseteq \mathbf{V}_H$ with $\theta_n - \theta \in \mathbf{E}$, $n \in \mathbb{N}$, and $\|\theta_n - \theta\|_{\mathbf{E}} \rightarrow 0$ for some fixed $\theta \in \mathbf{V}_H$, then we replace the phrase “w.r.t. \mathcal{S} ” by “at θ ” and “ $\dot{H}_{\mathcal{S}}$ ” by “ \dot{H}_{θ} ”.

(iii) If \mathcal{S} consists only of the constant sequence $\theta_n = \theta$, $n \in \mathbb{N}$, then we skip the phrase “uniformly” and replace the phrase “w.r.t. \mathcal{S} ” by “at θ ” and “ $\dot{H}_{\mathcal{S}}$ ” by “ \dot{H}_{θ} ”. In this case we may also replace “ $H(y_1) - H(y_2) \in \tilde{\mathbf{E}}$ for all $y_1, y_2 \in \mathbf{V}_H$ ” by “ $H(y) - H(\theta) \in \tilde{\mathbf{E}}$ for all $y \in \mathbf{V}_H$ ”.

(iv) If $\mathbf{E} = \mathbf{V}$, then we skip the phrase “quasi-”.

(v) If $\tilde{\mathbf{E}} = \tilde{\mathbf{V}}$, then we skip the phrase “with trace $\tilde{\mathbf{E}}$ ”.

The conventional notion of uniform Hadamard differentiability as used in Theorem 3.9.11 of [29] corresponds to the differentiability concept in (i) with \mathcal{S} as in (ii), \mathbf{E} as in (iv), and $\tilde{\mathbf{E}}$ as in (v). Proposition 4.1 below shows that it is beneficial to refrain from insisting on $\mathbf{E} = \mathbf{V}$ as in (iv). It was recently discussed in [3] that it can be also beneficial to refrain from insisting on the assumption of (ii). For $\mathbf{E} = \mathbf{V}$ (“non-quasi” case) uniform Hadamard differentiability in the sense of Definition B.1 in [3] corresponds to uniform Hadamard differentiability in the sense of our Definition 2.1 (part (i) and (iv)) when \mathcal{S} is chosen as the set of all sequences (θ_n) in a compact metric space $(\mathbf{K}_{\theta}, d_{\mathbf{K}})$ with $\theta \in \mathbf{K}_{\theta} \subseteq \mathbf{V}_H$ for which $d_{\mathbf{K}}(\theta_n, \theta) \rightarrow 0$. In Comment B.3 of [3] it is illustrated by means of the quantile functional that this notion of differentiability (subject to a suitable choice of $(\mathbf{K}_{\theta}, d_{\mathbf{K}})$) is strictly weaker than the notion of uniform Hadamard differentiability that was used in the classical delta-method for the almost sure bootstrap, Theorem 3.9.11 in [29]. Although this shows that the flexibility w.r.t. \mathcal{S} in our Definition 2.1 can be beneficial, it is somehow even more important that we allow for the “quasi” case.

Of course, the smaller the family \mathcal{S} the weaker the condition of uniform quasi-Hadamard differentiability w.r.t. \mathcal{S} . On the other hand, if the set \mathcal{S} is too small then condition (e) in Theorem 2.3 ahead may fail. That is, for an application of the functional delta-method in the form of Theorem 2.3 the set \mathcal{S} should be large enough for condition (e) to be fulfilled and small enough for being able to establish uniform quasi-Hadamard differentiability w.r.t. \mathcal{S} of the map H .

We now turn to the abstract delta-method. As mentioned in the introduction, convergence in distribution will always be considered for the open-ball σ -algebra. We will use the terminology *convergence in distribution*^o (symbolically \leadsto^o) for this sort of convergence; for details see the Appendix A and the Appendices A–C of [7]. In a separable metric space the notion of convergence in distribution^o boils down to the conventional notion of convergence in distribution for the Borel σ -algebra. In this case we use the symbol \leadsto instead of \leadsto^o .

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and (\hat{T}_n) be a sequence of maps

$$\hat{T}_n : \Omega \longrightarrow \mathbf{V}.$$

Regard $\omega \in \Omega$ as a sample drawn from \mathbb{P} , and $\widehat{T}_n(\omega)$ as a statistic derived from ω . Somewhat unconventionally, we do not (need to) require at this point that \widehat{T}_n is measurable w.r.t. any σ -algebra on \mathbf{V} . Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and set

$$(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}').$$

The probability measure \mathbb{P}' represents a random experiment that is run independently of the random sample mechanism \mathbb{P} . In the sequel, \widehat{T}_n will frequently be regarded as a map defined on the extension $\overline{\Omega}$ of Ω . Let

$$\widehat{T}_n^* : \overline{\Omega} \longrightarrow \mathbf{V}$$

be any map. Since $\widehat{T}_n^*(\omega, \omega')$ depends on both the original sample ω and the outcome ω' of the additional independent random experiment, we may regard \widehat{T}_n^* as a bootstrapped version of \widehat{T}_n . Moreover, let

$$\widehat{C}_n : \Omega \longrightarrow \mathbf{V}$$

be any map. As with \widehat{T}_n we often regard \widehat{C}_n as a map defined on the extension $\overline{\Omega}$ of Ω . We will use \widehat{C}_n together with a scaling sequence to get weak convergence results for \widehat{T}_n^* . The role of \widehat{C}_n is often played by \widehat{T}_n itself (cf. Example 3.3), but sometimes also by a different map (cf. Example 3.4). Assume that \widehat{T}_n , \widehat{T}_n^* , and \widehat{C}_n take values only in \mathbf{V}_H .

Let \mathcal{B}° and $\widetilde{\mathcal{B}}^\circ$ be the open-ball σ -algebras on \mathbf{E} and $\widetilde{\mathbf{E}}$ w.r.t. the norms $\|\cdot\|_{\mathbf{E}}$ and $\|\cdot\|_{\widetilde{\mathbf{E}}}$, respectively. Note that \mathcal{B}° coincides with the Borel σ -algebra on \mathbf{E} when $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ is separable. The same is true for $\widetilde{\mathcal{B}}^\circ$. Set $\widetilde{\widetilde{\mathbf{E}}} := \widetilde{\mathbf{E}} \times \widetilde{\mathbf{E}}$ and let $\widetilde{\widetilde{\mathcal{B}}}^\circ$ be the σ -algebra on $\widetilde{\widetilde{\mathbf{E}}}$ generated by the open balls w.r.t. the metric $\widetilde{\widetilde{d}}((\widetilde{x}_1, \widetilde{x}_2), (\widetilde{y}_1, \widetilde{y}_2)) := \max\{\|\widetilde{x}_1 - \widetilde{y}_1\|_{\widetilde{\mathbf{E}}}; \|\widetilde{x}_2 - \widetilde{y}_2\|_{\widetilde{\mathbf{E}}}\}$. Recall that $\widetilde{\widetilde{\mathcal{B}}}^\circ \subseteq \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ$, because any $\widetilde{\widetilde{d}}$ -open ball in $\widetilde{\widetilde{\mathbf{E}}}$ is the product of two $\|\cdot\|_{\widetilde{\mathbf{E}}}$ -open balls in $\widetilde{\mathbf{E}}$.

The following Theorem 2.2 is a consequence of Theorem A.4 in the Appendix A.2 as we assume that \widehat{T}_n takes values only in \mathbf{V}_H . The proof of the measurability statement of Theorem 2.2 is given in the proof of Theorem 2.3. Theorem 2.2 is stated here because, together with Theorem 2.3, it implies almost sure bootstrap consistency whenever the limit ξ is the same in Theorem 2.2 and Theorem 2.3.

Theorem 2.2 *Let (θ_n) be a sequence in \mathbf{V}_H and $\mathcal{S} := \{(\theta_n)\}$. Let $\mathbf{E}_0 \subseteq \mathbf{E}$ be a separable subspace and assume that $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let (a_n) be a sequence of positive real numbers with $a_n \rightarrow \infty$, and assume that the following assertions hold:*

(a) $a_n(\widehat{T}_n - \theta_n)$ takes values only in \mathbf{E} , is $(\mathcal{F}, \mathcal{B}^\circ)$ -measurable, and satisfies

$$a_n(\widehat{T}_n - \theta_n) \rightsquigarrow^\circ \xi \quad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}) \tag{4}$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.

- (b) $a_n(H(\widehat{T}_n) - H(\theta_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\mathcal{F}, \widetilde{\mathcal{B}}^\circ)$ -measurable.
- (c) H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$.

Then $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$ -measurable and

$$a_n(H(\widehat{T}_n) - H(\theta_n)) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) \quad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}).$$

Theorem 2.3 Let \mathcal{S} be any set of sequences in \mathbf{V}_H . Let $\mathbf{E}_0 \subseteq \mathbf{E}$ be a separable subspace and assume that $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let (a_n) be a sequence of positive real numbers with $a_n \rightarrow \infty$, and assume that the following assertions hold:

- (a) $a_n(\widehat{T}_n^* - \widehat{C}_n)$ takes values only in \mathbf{E} , is $(\overline{\mathcal{F}}, \mathcal{B}^\circ)$ -measurable, and satisfies

$$a_n(\widehat{T}_n^*(\omega, \cdot) - \widehat{C}_n(\omega)) \rightsquigarrow^\circ \xi \quad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}), \quad \mathbb{P}\text{-a.e. } \omega \quad (5)$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.

- (b) $a_n(H(\widehat{T}_n^*) - H(\widehat{C}_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\overline{\mathcal{F}}, \widetilde{\mathcal{B}}^\circ)$ -measurable.
- (c) H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$.
- (d) The uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$ can be extended from \mathbf{E}_0 to \mathbf{E} such that the extension $\dot{H}_{\mathcal{S}} : \mathbf{E} \rightarrow \widetilde{\mathbf{E}}$ is $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable and continuous at every point of \mathbf{E}_0 .
- (e) $(\widehat{C}_n(\omega)) \in \mathcal{S}$ for \mathbb{P} -a.e. ω .
- (f) The map $h : \widetilde{\mathbf{E}} \rightarrow \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\widetilde{\mathcal{B}}^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable.

Then $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$ -measurable and

$$a_n(H(\widehat{T}_n^*(\omega, \cdot)) - H(\widehat{C}_n(\omega))) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) \quad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}), \quad \mathbb{P}\text{-a.e. } \omega. \quad (6)$$

Remark 2.4 In condition (a) of Theorem 2.3 it is assumed that $a_n(\widehat{T}_n^* - \widehat{C}_n)$ is $(\overline{\mathcal{F}}, \mathcal{B}^\circ)$ -measurable for $\overline{\mathcal{F}} := \mathcal{F} \otimes \mathcal{F}'$. Thus the mapping $\omega' \mapsto a_n(\widehat{T}_n^*(\omega, \omega') - \widehat{C}_n(\omega))$ is $(\mathcal{F}', \mathcal{B}^\circ)$ -measurable for every fixed $\omega \in \Omega$. That is, $a_n(\widehat{T}_n^*(\omega, \cdot) - \widehat{C}_n(\omega))$ can be seen as an $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable on $(\Omega', \mathcal{F}', \mathbb{P}')$ for every fixed $\omega \in \Omega$, so that assertion (5) makes sense. By the same line of reasoning one can regard $a_n(H(\widehat{T}_n^*(\omega, \cdot)) - H(\widehat{C}_n(\omega)))$ as an $(\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ)$ -valued random variable on $(\Omega', \mathcal{F}', \mathbb{P}')$ for every fixed $\omega \in \Omega$, so that also assertion (6) makes sense. \diamond

Remark 2.5 Condition (c) in Theorem 2.2 (resp. Theorem 2.3) assumes that the trace is given by $\tilde{\mathbf{E}}$, which implies that the first part of condition (b) in Theorem 2.2 (resp. Theorem 2.3) is automatically satisfied. \diamond

Remark 2.6 Condition (f) of Theorem 2.3 is automatically fulfilled when $(\tilde{\mathbf{E}}, \|\cdot\|_{\tilde{\mathbf{E}}})$ is separable. Indeed, in this case we have $\overline{\tilde{\mathcal{B}}^\circ} = \tilde{\mathcal{B}}^\circ \otimes \tilde{\mathcal{B}}^\circ$ so that every continuous map $h : \overline{\tilde{\mathbf{E}}} \rightarrow \tilde{\mathbf{E}}$ (such as $h(\tilde{x}_1, \tilde{x}_2) := \tilde{x}_1 - \tilde{x}_2$) is $(\overline{\tilde{\mathcal{B}}^\circ}, \tilde{\mathcal{B}}^\circ)$ -measurable. \diamond

Proof of Theorem 2.3 First note that by the assumption imposed on ξ (cf. assumption (a)) and assumption (c) the map $\dot{H}_S(\xi)$ is $(\mathcal{F}_0, \tilde{\mathcal{B}}^\circ)$ -measurable. Next note that

$$\begin{aligned} & a_n(H(\hat{T}_n^*(\omega, \omega')) - H(\hat{C}_n(\omega))) \\ &= \{a_n(H(\hat{T}_n^*(\omega, \omega')) - H(\hat{C}_n(\omega))) - \dot{H}_S(a_n(\hat{T}_n^*(\omega, \omega') - \hat{C}_n(\omega)))\} \\ & \quad + \dot{H}_S(a_n(\hat{T}_n^*(\omega, \omega') - \hat{C}_n(\omega))) \\ &=: S_1(\omega, \omega') + S_2(\omega, \omega'). \end{aligned}$$

By (5) in assumption (a) and the Continuous Mapping theorem in the form of [8, Theorem 6.4] (along with $\mathbb{P}_0 \circ \xi^{-1}[\mathbf{E}_0] = 1$ and the continuity of \dot{H}_S), we have that $S_2(\omega, \cdot) \xrightarrow{\circ} \dot{H}_S(\xi)$ for \mathbb{P} -a.e. ω . Moreover, for every fixed ω we have that $\omega' \mapsto S_1(\omega, \omega')$ is $(\mathcal{F}', \tilde{\mathcal{B}}^\circ)$ -measurable by assumption (f), and for \mathbb{P} -a.e. ω we have

$$a_n(H_n(\hat{T}_n^*(\omega, \cdot)) - H_n(\hat{C}_n(\omega))) - \dot{H}_S(a_n(\hat{T}_n^*(\omega, \omega') - \hat{C}_n(\omega))) \xrightarrow{\mathbb{P}^\circ} 0_{\tilde{\mathbf{E}}}$$

by part (ii) of Theorem A.4 (recall that \hat{T}_n^* was assumed to take values only in \mathbf{V}_H), where $\rightarrow^{\mathbb{P}, \circ}$ refers to convergence in probability^o (cf. Section A.1) and $\hat{T}_n^*(\omega, \cdot)$, $\hat{C}_n(\omega)$, $\{(\hat{C}_n(\omega))\}$ play the roles of $\hat{T}_n(\cdot)$, θ_n , \mathcal{S} , respectively. Hence, from Corollary A.3 we get that (6) holds. \square

3. Application to plug-in estimators of statistical functionals

Let \mathbf{D} be the space of all càdlàg functions v on \mathbb{R} with finite sup-norm $\|v\|_\infty := \sup_{t \in \mathbb{R}} |v(t)|$, and \mathcal{D} be the σ -algebra on \mathbf{D} generated by the one-dimensional coordinate projections π_t , $t \in \mathbb{R}$, given by $\pi_t(v) := v(t)$. Let $\phi : \mathbb{R} \rightarrow [1, \infty)$ be a weight function, i.e. a continuous function being non-increasing on $(-\infty, 0]$ and non-decreasing on $[0, \infty)$. Let \mathbf{D}_ϕ be the subspace of \mathbf{D} consisting of all $x \in \mathbf{D}$ satisfying $\|x\|_\phi := \|x\phi\|_\infty < \infty$ and $\lim_{|t| \rightarrow \infty} |x(t)| = 0$. The latter condition automatically holds when $\lim_{|t| \rightarrow \infty} \phi(t) = \infty$. Let $\mathcal{D}_\phi := \mathcal{D} \cap \mathbf{D}_\phi$ be the trace σ -algebra on \mathbf{D}_ϕ . The σ -algebra on \mathbf{D}_ϕ generated by the

$\|\cdot\|_\phi$ -open balls will be denoted by \mathcal{B}_ϕ° . Lemma 4.1 in [7] shows that it coincides with \mathcal{D}_ϕ .

Let $\mathbf{C}_\phi \subseteq \mathbf{D}_\phi$ be a $\|\cdot\|_\phi$ -separable subspace and assume $\mathbf{C}_\phi \in \mathcal{D}_\phi$. Moreover, let $H : \mathbf{D}(H) \rightarrow \tilde{\mathbf{V}}$ be a map defined on a set $\mathbf{D}(H)$ of distribution functions of finite (not necessarily probability) Borel measures on \mathbb{R} , where $\tilde{\mathbf{V}}$ is any vector space. In particular, $\mathbf{D}(H) \subseteq \mathbf{D}$. In the following, \mathbf{D} , $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi)$, \mathbf{C}_ϕ , and $\mathbf{D}(H)$ will play the roles of \mathbf{V} , $(\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_\mathbf{E})$, \mathbf{E}_0 , and \mathbf{V}_H , respectively. As before we let $(\tilde{\mathbf{E}}, \|\cdot\|_{\tilde{\mathbf{E}}})$ be a normed subspace of $\tilde{\mathbf{V}}$ equipped with the corresponding open-ball σ -algebra $\tilde{\mathcal{B}}^\circ$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(F_n) \subseteq \mathbf{D}(H)$ be any sequence and (X_i) be a sequence of real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover let $\hat{F}_n : \Omega \rightarrow \mathbf{D}$ be the empirical distribution function of X_1, \dots, X_n , which will play the role of \hat{T}_n . It is defined by

$$\hat{F}_n := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{[X_i, \infty)}. \quad (7)$$

Assume that \hat{F}_n takes values only in $\mathbf{D}(H)$. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and set $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$. Moreover, let $\hat{F}_n^* : \bar{\Omega} \rightarrow \mathbf{D}$ be any map. Assume that \hat{F}_n^* take values only in $\mathbf{D}(H)$. Furthermore, let $\hat{C}_n : \Omega \rightarrow \mathbf{D}$ be any map that takes values only in $\mathbf{D}(H)$. In the present setting Theorems 2.2 and 2.3 can be reformulated as follows, where we recall from Remark 2.6 that condition (f) of Theorem 2.3 is automatically fulfilled when $(\tilde{\mathbf{E}}, \|\cdot\|_{\tilde{\mathbf{E}}})$ is separable.

Corollary 3.1 *Let (F_n) be a sequence in $\mathbf{D}(H)$ and $\mathcal{S} := \{(F_n)\}$. Let (a_n) be a sequence of positive real numbers with $a_n \rightarrow \infty$, and assume that the following assertions hold:*

(a) $a_n(\hat{F}_n - F_n)$ takes values only in \mathbf{D}_ϕ and satisfies

$$a_n(\hat{F}_n - F_n) \rightsquigarrow^\circ B \quad \text{in } (\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi) \quad (8)$$

for some $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ)$ -valued random variable B on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $B(\Omega_0) \subseteq \mathbf{C}_\phi$.

(b) $a_n(H(\hat{F}_n) - H(F_n))$ takes values only in $\tilde{\mathbf{E}}$ and is $(\mathcal{F}, \tilde{\mathcal{B}}^\circ)$ -measurable.

(c) H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{C}_\phi \langle \mathbf{D}_\phi \rangle$ with trace $\tilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$.

Then $\dot{H}_\mathcal{S}(B)$ is $(\mathcal{F}_0, \tilde{\mathcal{B}}^\circ)$ -measurable and

$$a_n(H(\hat{F}_n) - H(F_n)) \rightsquigarrow^\circ \dot{H}_\mathcal{S}(B) \quad \text{in } (\tilde{\mathbf{E}}, \tilde{\mathcal{B}}^\circ, \|\cdot\|_{\tilde{\mathbf{E}}}).$$

Note that the measurability assumption in condition (a) of Theorem 2.2 is automatically satisfied in the present setting (and is therefore omitted in condition (a) of Corollary 3.1). Indeed, $a_n(\hat{F}_n - F)$ is $(\mathcal{F}, \mathcal{B}_\phi^\circ)$ -measurable, because it is easily seen to be $(\mathcal{F}, \mathcal{D}_\phi)$ -measurable and we have noted above that $\mathcal{B}_\phi^\circ = \mathcal{D}_\phi$.

Corollary 3.2 *Let \mathcal{S} be any set of sequences in $\mathbf{D}(H)$. Let (a_n) be a sequence of positive real numbers with $a_n \rightarrow \infty$, and assume that the following assertions hold:*

(a) $a_n(\widehat{F}_n^* - \widehat{C}_n)$ takes values only in \mathbf{D}_ϕ , is $(\overline{\mathcal{F}}, \mathcal{B}_\phi^\circ)$ -measurable, and

$$a_n(\widehat{F}_n^*(\omega, \cdot) - \widehat{C}_n(\omega)) \rightsquigarrow^\circ B \quad \text{in } (\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi), \quad \mathbb{P}\text{-a.e. } \omega \quad (9)$$

for some $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ)$ -valued random variable B on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $B(\Omega_0) \subseteq \mathbf{C}_\phi$.

(b) $a_n(H(\widehat{F}_n^*) - H(\widehat{C}_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\overline{\mathcal{F}}, \widetilde{\mathcal{B}}^\circ)$ -measurable.

(c) H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{C}_\phi \langle \mathbf{D}_\phi \rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$.

(d) The uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$ can be extended from \mathbf{C}_ϕ to \mathbf{D}_ϕ such that the extension $\dot{H}_\mathcal{S} : \mathbf{D}_\phi \rightarrow \widetilde{\mathbf{E}}$ is $(\mathcal{B}_\phi^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable, and continuous at every point of \mathbf{C}_ϕ .

(e) $(\widehat{C}_n(\omega)) \in \mathcal{S}$ for \mathbb{P} -a.e. ω .

(f) The map $h : \overline{\widetilde{\mathbf{E}}} \rightarrow \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\overline{\widetilde{\mathcal{B}}^\circ}, \widetilde{\mathcal{B}}^\circ)$ -measurable.

Then $\dot{H}_\mathcal{S}(B)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$ -measurable and

$$a_n(H(\widehat{F}_n^*(\omega, \cdot)) - H(\widehat{C}_n(\omega))) \rightsquigarrow^\circ \dot{H}_\mathcal{S}(B) \quad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}), \quad \mathbb{P}\text{-a.e. } \omega.$$

The following two examples illustrate \widehat{F}_n^* and \widehat{C}_n . In S1. and S2. in the first example, i.e. Example 3.3, we have $\widehat{C}_n = \widehat{F}_n$, and in S3. of this example as well as in the second example, i.e. Example 3.4, \widehat{C}_n may differ from \widehat{F}_n . Examples for uniformly quasi-Hadamard differentiable functionals H can be found in Section 4. In the examples in Sections 4.1 and 4.3 we have $\widetilde{\mathbf{V}} = \widetilde{\mathbf{E}} = \mathbb{R}$, and in the Example in Section 4.2 we have $\widetilde{\mathbf{V}} = \mathbf{D}$ and $\widetilde{\mathbf{E}} = \mathbf{D}_\phi$ for some ϕ .

Example 3.3 Let (X_i) be a sequence of i.i.d. real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F , and \widehat{F}_n be given by (7). Let (W_{ni}) be a triangular array of nonnegative real-valued random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ such that (W_{n1}, \dots, W_{nn}) is an exchangeable random vector for every $n \in \mathbb{N}$, and define the map $\widehat{F}_n^* : \overline{\Omega} \rightarrow \mathbf{D}$ by

$$\widehat{F}_n^*(\omega, \omega') := \frac{1}{n} \sum_{i=1}^n W_{ni}(\omega') \mathbb{1}_{[X_i(\omega), \infty)}. \quad (10)$$

Note that the sequence (X_i) and the triangular array (W_{ni}) regarded as families of random variables on the product space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ are independent. Of course, we will tacitly assume that $(\Omega', \mathcal{F}', \mathbb{P}')$ is rich enough to host all

the random variables used below. Similar as in Lemma 5.1 of [7] it can be shown that $a_n(\hat{F}_n^* - \hat{C}_n)$, with $\hat{C}_n := \overline{W}_n \hat{F}_n$, takes values only in \mathbf{D}_ϕ and is $(\overline{\mathcal{F}}, \mathcal{D}_\phi)$ -measurable, where $\overline{W}_n := \frac{1}{n} \sum_{i=1}^n W_{ni}$. That is, the first part of condition (a) of Corollary 3.2 holds true for $\hat{C}_n := \overline{W}_n \hat{F}_n$. Now assume that F satisfies $\int \phi^2 dF < \infty$ and that the following three assertions hold.

- A1. $\sup_{n \in \mathbb{N}} \int_0^\infty \mathbb{P}'[|W_{n1} - \overline{W}_n| > t]^{1/2} dt < \infty$.
- A2. $\frac{1}{\sqrt{n}} \mathbb{E}'[\max_{1 \leq i \leq n} |W_{ni} - \overline{W}_n|] \rightarrow 0$.
- A3. $\frac{1}{n} \sum_{i=1}^n (W_{ni} - \overline{W}_n)^2 \rightarrow 1$ in \mathbb{P}' -probability.

Then, arguing as in Example 4.3 and Section 5.1 of [7], results in [28] and [29] imply that respectively condition (a) of Corollary 3.1 (with $F_n := F$) and the second part of condition (a) of Corollary 3.2 (with $\hat{C}_n := \overline{W}_n \hat{F}_n$) hold for $a_n := \sqrt{n}$ and $B := B_F$, where B_F is an F -Brownian bridge, i.e. a centered Gaussian process with covariance function $\Gamma(t_0, t_1) = F(t_0 \wedge t_1) \overline{F}(t_0 \vee t_1)$. Here \mathbf{C}_ϕ can be chosen to be the set $\mathbf{C}_{\phi, F}$ of all $v \in \mathbf{D}_\phi$ whose discontinuities are also discontinuities of F .

Examples 3.6.9, 3.6.10, and 3.6.12 in [29] show that conditions A1.–A3. are satisfied if one of the following three specific settings is met:

- S1. The random vector (W_{n1}, \dots, W_{nn}) is multinomially distributed according to the parameters n and $p_1 = \dots = p_n = \frac{1}{n}$ for every $n \in \mathbb{N}$.
- S2. $W_{ni} = Y_i / \overline{Y}_n$ for every $i = 1, \dots, n$ and $n \in \mathbb{N}$, where $\overline{Y}_n := \frac{1}{n} \sum_{j=1}^n Y_j$ and (Y_j) is any sequence of nonnegative i.i.d. random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ with $\int_0^\infty \mathbb{P}'[Y_1 > t]^{1/2} dt < \infty$ and $\mathbb{V}\text{ar}'[Y_1]^{1/2} = \mathbb{E}'[Y_1] > 0$.
- S3. $W_{ni} = Y_i$ for every $i = 1, \dots, n$ and $n \in \mathbb{N}$, where (Y_i) is any sequence of non-negative i.i.d. random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ with $\int_0^\infty \mathbb{P}'[Y_1 > t]^{1/2} dt < \infty$ and $\mathbb{V}\text{ar}'[Y_1] = 1$.

Setting S1. is nothing but *Efron's bootstrap* [14] and Setting S3. is sometimes referred to as *wild bootstrap*. If in Setting S2. the distribution of Z_1 is the exponential distribution with parameter 1, then the resulting scheme is in line with the *Bayesian bootstrap* of Rubin [26]. Note that in Settings S1. and S2. we have $\overline{W}_n = 1$ and thus $\hat{C}_n = \hat{F}_n$. This implies that condition (e) holds if \mathcal{S} is (any subset of) the set of all sequences (G_n) of distribution functions on \mathbb{R} satisfying $G_n - F \in \mathbf{D}_\phi$, $n \in \mathbb{N}$, and $\|G_n - F\|_\phi \rightarrow 0$; see, for instance, Theorem 2.1 in [31]. \diamond

Example 3.4 Let (X_i) be a strictly stationary sequence of β -mixing random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F , and \hat{F}_n be given by (7). Let (ℓ_n) be a sequence of integers such that $\ell_n \nearrow \infty$ as $n \rightarrow \infty$, and $\ell_n < n$ for all $n \in \mathbb{N}$. Set $k_n := \lceil n/\ell_n \rceil$ for all

$n \in \mathbb{N}$. Let $(I_{nj})_{n \in \mathbb{N}, 1 \leq j \leq k_n}$ be a triangular array of random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ such that I_{n1}, \dots, I_{nk_n} are i.i.d. according to the uniform distribution on $\{1, \dots, n - \ell_n + 1\}$ for every $n \in \mathbb{N}$. Define the map $\widehat{F}_n^* : \overline{\Omega} \rightarrow \mathbf{D}$ by (10) with

$$W_{ni}(\omega') := \sum_{j=1}^{k_n-1} \mathbb{1}_{\{I_{nj} \leq i \leq I_{nj} + \ell_n - 1\}}(\omega') + \mathbb{1}_{\{I_{nk_n} \leq i \leq I_{nk_n} + (n - (k_n - 1)\ell_n) - 1\}}(\omega'). \quad (11)$$

Note that, as before, the sequence (X_i) and the triangular array (W_{ni}) regarded as families of random variables on the product space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ are independent. At an informal level this means that given a sample X_1, \dots, X_n , we pick $k_n - 1$ blocks of length ℓ_n and one block of length $n - (k_n - 1)\ell_n$ in the sample X_1, \dots, X_n , where the start indices $I_{n1}, I_{n2}, \dots, I_{nk_n}$ are chosen independently and uniformly in the set of indices $\{1, \dots, n - \ell_n + 1\}$:

$$\begin{aligned} \text{block 1:} & \quad X_{I_{n1}}, X_{I_{n1}+1}, \dots, X_{I_{n1}+\ell_n-1} \\ \text{block 2:} & \quad X_{I_{n2}}, X_{I_{n2}+1}, \dots, X_{I_{n2}+\ell_n-1} \\ & \quad \vdots \\ \text{block } k_n - 1: & \quad X_{I_{n(k_n-1)}}, X_{I_{n(k_n-1)}+1}, \dots, X_{I_{n(k_n-1)}+\ell_n-1} \\ \text{block } k_n: & \quad X_{I_{nk_n}}, X_{I_{nk_n}+1}, \dots, X_{I_{nk_n}+(n-(k_n-1)\ell_n)-1}. \end{aligned}$$

The bootstrapped empirical distribution function \widehat{F}_n^* is then defined to be the distribution function of the discrete finite (not necessarily probability) measure with atoms X_1, \dots, X_n carrying masses W_{n1}, \dots, W_{nn} respectively, where W_{ni} specifies the number of blocks which contain X_i . Similar as in Lemma 5.3 in [7] it follows that $a_n(\widehat{F}_n^* - \widehat{C}_n)$, with $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*]$, takes values only in \mathbf{D}_ϕ and is $(\overline{\mathcal{F}}, \mathcal{D}_\phi)$ -measurable. That is, the first part of condition (a) of Corollary 3.2 holds true for $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*]$. Now assume that the following assertions hold:

- A1. $\int \phi^p dF < \infty$ for some $p > 4$.
- A2. The sequence of random variables (X_i) is strictly stationary and β -mixing with mixing coefficients (β_i) satisfying $\beta_i \leq c\delta^i$ for some constants $c > 0$ and $\delta \in (0, 1)$.
- A3. The block length ℓ_n satisfies $\ell_n = \mathcal{O}(n^\gamma)$ for some $\gamma \in (0, 1/2)$.

Then, as discussed in Example 4.4 and Section 5.2 of [7], it can be derived from a result in [1] that under assumptions A1. and A2. we have that condition (a) of Corollary 3.1 holds for $a_n := \sqrt{n}$, $B := B_F$, and $F_n := F$, where B_F is a centered Gaussian process with covariance function $\Gamma(t_0, t_1) = F(t_0 \wedge t_1)(1 - F(t_0 \vee t_1)) + \sum_{i=0}^1 \sum_{k=2}^\infty \text{Cov}(\mathbb{1}_{\{X_i \leq t_i\}}, \mathbb{1}_{\{X_k \leq t_{1-i}\}})$. Here \mathbf{C}_ϕ can be chosen to be the set $\mathbf{C}_{\phi, F}$ of all $v \in \mathbf{D}_\phi$ whose discontinuities are also discontinuities of F . Moreover, Theorem 3.5 below

shows that under the assumptions A1.–A3. the second part of condition (a) (i.e. (9)) and condition (e) of Corollary 3.2 hold for

$$\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*] = \frac{1}{n} \sum_{i=1}^n w_{ni} \mathbb{1}_{[X_i, \infty)} \quad \text{with} \quad w_{ni} := \mathbb{E}'[W_{ni}] \quad (12)$$

and the same choice of a_n , B , and F_n , when \mathcal{S} is the set of all sequences $(G_n) \subseteq \mathbf{D}(H)$ with $G_n - F \in \mathbf{D}_\phi$, $n \in \mathbb{N}$, and $\|G_n - F\|_\phi \rightarrow 0$. Note that

$$w_{ni} = \begin{cases} k_n \frac{i}{n-\ell_n+1} & , \quad i = 1, \dots, n - (k_n - 1)\ell_n \\ (k_n - 1) \frac{i}{n-\ell_n+1} + \frac{n-(k_n-1)\ell_n}{n-\ell_n+1} & , \quad i = n - (k_n - 1)\ell_n + 1, \dots, \ell_n \\ (k_n - 1) \frac{\ell_n}{n-\ell_n+1} + \frac{n-(k_n-1)\ell_n}{n-\ell_n+1} = \frac{n}{n-\ell_n+1} & , \quad i = \ell_n + 1, \dots, n - \ell_n \\ (k_n - 1) \frac{n-i+1}{n-\ell_n+1} + \frac{2n-k_n\ell_n-i+1}{n-\ell_n+1} & , \quad i = n - \ell_n + 1, \dots, n - (k_n\ell_n - n) \\ (k_n - 1) \frac{n-i+1}{n-\ell_n+1} & , \quad i = n - (k_n\ell_n - n) + 1, \dots, n \end{cases} \quad (13)$$

which can be verified easily. \diamond

Further examples for condition (a) in Corollary 3.2 for dependent observations can, for example, be found in [9, 21, 22].

Theorem 3.5 *In the setting of Example 3.4 assume that assertions A1.–A3. hold, and let \mathcal{S} be the set of all sequences $(G_n) \subseteq \mathbf{D}(H)$ with $G_n - F \in \mathbf{D}_\phi$, $n \in \mathbb{N}$, and $\|G_n - F\|_\phi \rightarrow 0$. Then the second part of assertion (a) (i.e. (9)) and assertion (e) in Corollary 3.2 hold.*

Proof *Proof of second part of (a):* It is enough to show that under assumptions A1.–A3. the assumptions (A1)–(A4) of Theorem 1 in [10] hold when the class of functions is $\mathbf{F}_\phi := \mathbf{F}_\phi^- \cup \mathbf{F}_\phi^+$, where $\mathbf{F}_\phi^- := \{f_x : x \leq 0\}$ and $\mathbf{F}_\phi^+ := \{f_x : x > 0\}$ with $f_x(\cdot) := \phi(x) \mathbb{1}_{(-\infty, x]}(\cdot)$ for $x \leq 0$ and $f_x(\cdot) := -\phi(x) \mathbb{1}_{(x, \infty)}(\cdot)$ for $x > 0$. Due to A2. and A3. we only have to verify assumptions (A3) and (A4) of Theorem 1 in [10]. That is, we will show that the following two assertions hold.

- 1) There exist constants $b, c > 0$ such that $N_{[\cdot]}(\varepsilon, \mathbf{F}_\phi, \|\cdot\|_p) \leq c\varepsilon^{-b}$ for all $\varepsilon > 0$.
- 2) $\int \bar{f}^p dF < \infty$ for the envelope function $\bar{f}(z) := \sup_{x \in \mathbb{R}} |f_x(z)|$.

Here the bracketing number $N_{[\cdot]}(\varepsilon, \mathbf{F}_\phi, \|\cdot\|_p)$ is the minimal number of ε -brackets w.r.t. $\|\cdot\|_p$ (L^p -norm w.r.t. dF) to cover \mathbf{F}_ϕ , where an ε -bracket w.r.t. $\|\cdot\|_p$ is the set, $[\ell, u]$, of all functions f with $\ell \leq f \leq u$ for some Borel measurable functions $\ell, u : \mathbb{R} \rightarrow \mathbb{R}_+$ with $\ell \leq u$ pointwise and $\|u - \ell\|_p \leq \varepsilon$.

1): We will only show that 1) with \mathbf{F}_ϕ replaced by \mathbf{F}_ϕ^- holds true. Analogously one can show that the same holds true for \mathbf{F}_ϕ^+ (and therefore for \mathbf{F}_ϕ). On the one hand,

since $I_p^- := \int_{(-\infty, 0]} \phi^p dF < \infty$ by assumption (a), we can find for every $\varepsilon > 0$ a finite partition $-\infty = y_0^\varepsilon < y_1^\varepsilon < \dots < y_{k_\varepsilon}^\varepsilon = 0$ such that

$$\max_{i=1, \dots, k_\varepsilon} \int_{(y_{i-1}^\varepsilon, y_i^\varepsilon]} \phi^p dF \leq (\varepsilon/2)^p \quad (14)$$

and $k_\varepsilon \leq \lceil I_p^- / (\varepsilon/2)^p \rceil$. On the other hand, using integration by parts we obtain

$$\int_{(-\infty, 0]} F d(-\phi^p) = \phi(0)F(0) - \int_{(-\infty, 0]} (-\phi^p) dF = \phi(0)F(0) + I_p^-,$$

so that we can find a finite partition $-\infty = z_0^\varepsilon < z_1^\varepsilon < \dots < z_{m_\varepsilon}^\varepsilon = 0$ such that

$$\max_{i=1, \dots, m_\varepsilon} \int_{(z_{i-1}^\varepsilon, z_i^\varepsilon]} F d(-\phi^p) \leq (\varepsilon/2)^p \quad (15)$$

and $m_\varepsilon \leq \lceil (\phi(0)F(0) + I_p^-) / (\varepsilon/2)^p \rceil$.

Now let $-\infty = x_0^\varepsilon < x_1^\varepsilon < \dots < x_{k_\varepsilon + m_\varepsilon}^\varepsilon = 0$ be the partition consisting of all points y_i^ε and z_i^ε , and set

$$\begin{aligned} \ell_i^\varepsilon(\cdot) &:= \phi(x_i^\varepsilon) \mathbb{1}_{(-\infty, x_{i-1}^\varepsilon]}(\cdot), \\ u_i^\varepsilon(\cdot) &:= \phi(x_{i-1}^\varepsilon) \mathbb{1}_{(-\infty, x_{i-1}^\varepsilon]}(\cdot) + \phi(\cdot) \mathbb{1}_{(x_{i-1}^\varepsilon, x_i^\varepsilon]}(\cdot). \end{aligned} \quad (16)$$

Then $\ell_i^\varepsilon \leq u_i^\varepsilon$. Moreover

$$\begin{aligned} \|u_i^\varepsilon - \ell_i^\varepsilon\|_p &= \left(\int (u_i^\varepsilon - \ell_i^\varepsilon)^p dF \right)^{1/p} \\ &\leq \left(\int_{(-\infty, x_{i-1}^\varepsilon]} (\phi(x_{i-1}^\varepsilon) - \phi(x_i^\varepsilon))^p dF \right)^{1/p} + \left(\int_{(x_{i-1}^\varepsilon, x_i^\varepsilon]} \phi^p dF \right)^{1/p} \\ &\leq \left(\int_{(-\infty, x_{i-1}^\varepsilon]} (\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) dF \right)^{1/p} + \varepsilon/2 \\ &\leq \left((\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) F(x_{i-1}^\varepsilon) \right)^{1/p} + \varepsilon/2 \end{aligned}$$

where we used Minkovski's inequality and (14), and that ϕ is non-increasing on $(-\infty, 0]$ and $x_{i-1}^\varepsilon \leq x_i^\varepsilon$. Since F is at least $F(x_{i-1}^\varepsilon)$ on $(x_{i-1}^\varepsilon, x_i^\varepsilon]$, we have

$$(\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) F(x_{i-1}^\varepsilon) \leq \int_{(x_{i-1}^\varepsilon, x_i^\varepsilon]} F d(-\phi^p) \leq (\varepsilon/2)^p$$

due to (15). Thus $\|u_i^\varepsilon - \ell_i^\varepsilon\|_p \leq \varepsilon$, so that $[\ell_i^\varepsilon, u_i^\varepsilon]$ provides an ε -bracket w.r.t. $\|\cdot\|_p$. It is moreover obvious that the ε -brackets $[\ell_i^\varepsilon, u_i^\varepsilon]$, $i = 1, \dots, k_\varepsilon + m_\varepsilon$, cover \mathbf{F}_ϕ^- . Thus, $N_{[\cdot]}(\varepsilon, \mathbf{F}_\phi^-, \|\cdot\|_p) \leq c\varepsilon^{-p}$ for a suitable constant $c > 0$ and all $\varepsilon > 0$.

2): The envelope function \bar{f} is given by $\bar{f}(y) = \phi(y)$ for $y \leq 0$ and by $\bar{f}(y) = \phi(y-)$ for $y > 0$ (recall that ϕ is continuous) for $y > 0$. Then under assumption (a) the integrability condition 2) holds.

Proof of (e): We have to show that $\|\widehat{C}_n - F\|_\phi = \sup_{x \in \mathbb{R}} |\widehat{C}_n(x) - F(x)|\phi(x) \rightarrow 0$ \mathbb{P} -a.s. We will only show that

$$\sup_{x \in (-\infty, 0]} |\widehat{C}_n(x) - F(x)|\phi(x) \longrightarrow 0 \quad \mathbb{P}\text{-a.s.}, \quad (17)$$

because the analogue for the positive real line can be shown in the same way. Let ℓ_i^ε and u_i^ε be as defined in (16). By assumption A1. we have $\int \phi dF < \infty$, so that similar as above we can find a finite partition $-\infty = x_0^\varepsilon < x_1^\varepsilon < \dots < x_{k_\varepsilon + m_\varepsilon}^\varepsilon = 0$ such that $[\ell_i^\varepsilon, u_i^\varepsilon]$, $i = 1, \dots, k_\varepsilon + m_\varepsilon$, are ε -brackets w.r.t. $\|\cdot\|_1$ (L^1 -norm w.r.t. F) covering the class $\mathbf{F}_\phi := \{f_x : x \in \mathbb{R}\}$ introduced above. We will proceed in two steps.

Step 1. First we will show that

$$\sup_{x \leq 0} |\widehat{C}_n(x) - F(x)|\phi(x) \leq \max_{i=1, \dots, k_\varepsilon + m_\varepsilon} \max \left\{ \int u_i^\varepsilon d(\widehat{C}_n - F); \int \ell_i^\varepsilon d(F - \widehat{C}_n) \right\} + \varepsilon \quad (18)$$

holds true for every $\varepsilon > 0$. Since $(\widehat{C}_n(x) - F(x))\phi(x) = \int f_x d\widehat{C}_n - \int f_x dF$, for (18) it suffices to show

$$\begin{aligned} \sup_{x \leq 0} \left| \int f_x d\widehat{C}_n - \int f_x dF \right| \\ \leq \max_{i=1, \dots, k_\varepsilon + m_\varepsilon} \max \left\{ \int u_i^\varepsilon d(\widehat{C}_n - F); \int \ell_i^\varepsilon d(F - \widehat{C}_n) \right\} + \varepsilon. \end{aligned} \quad (19)$$

To prove (19), we note that for every $x \in (-\infty, y]$ there is some $i_x \in \{1, \dots, k_\varepsilon + m_\varepsilon\}$ such that $f_x \in [\ell_{i_x}^\varepsilon, u_{i_x}^\varepsilon]$; cf. Step 1. Therefore, since $[\ell_{i_x}^\varepsilon, u_{i_x}^\varepsilon]$ is an ε -bracket w.r.t. $\|\cdot\|_1$,

$$\begin{aligned} \int f_x d\widehat{C}_n - \int f_x dF &\leq \int u_{i_x}^\varepsilon d\widehat{C}_n - \int f_x dF \\ &= \int u_{i_x}^\varepsilon d(\widehat{C}_n - F) + \int (u_{i_x}^\varepsilon - f_x) dF \\ &\leq \int u_{i_x}^\varepsilon d(\widehat{C}_n - F) + \int (u_{i_x}^\varepsilon - \ell_{i_x}^\varepsilon) dF \\ &\leq \max_{i=1, \dots, k_\varepsilon + m_\varepsilon} \int u_i^\varepsilon d(\widehat{C}_n - F) + \varepsilon. \end{aligned}$$

Analogously we obtain

$$\int f_x d\widehat{C}_n - \int f_x dF \geq - \left(\max_{i=1, \dots, k_\varepsilon + m_\varepsilon} \int \ell_i^\varepsilon d(F - \widehat{C}_n) + \varepsilon \right).$$

That is, (18) holds true.

Step 2. Because of (18), for (17) to be true it suffices to show that

$$\int \ell_i^\varepsilon d(F - \widehat{C}_n) \longrightarrow 0 \quad \text{and} \quad \int u_i^\varepsilon d(\widehat{C}_n - F) \longrightarrow 0 \quad \mathbb{P}\text{-a.s.} \quad (20)$$

for every $i = 1, \dots, k_\varepsilon + m_\varepsilon$. We will only show the second convergence in (20), the first convergence can be shown even easier. We have

$$\begin{aligned} \int u_i^\varepsilon d(\widehat{C}_n - F) &= \frac{1}{n} \sum_{j=1}^n \left(w_{ni} \phi(y_{i-1}^\varepsilon) \mathbb{1}_{(-\infty, y_{i-1}^\varepsilon]}(X_j) - \mathbb{E}[\phi(y_{i-1}^\varepsilon) \mathbb{1}_{(-\infty, y_{i-1}^\varepsilon]}(X_1)] \right) \\ &\quad + \frac{1}{n} \sum_{j=1}^n \left(w_{ni} \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) - \mathbb{E}[\phi(X_1) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_1)] \right) \\ &=: S_1(n) + S_2(n). \end{aligned}$$

The first summand on the right-hand side of

$$\begin{aligned} S_2(n) &= \frac{1}{n} \sum_{j=1}^n \left(\phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) - \mathbb{E}[\phi(X_1) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_1)] \right) \\ &\quad + \frac{1}{n} \sum_{j=1}^n (w_{ni} - 1) \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) \end{aligned}$$

converges \mathbb{P} -a.s. to 0 by Theorem 1 (ii) (and Application 5, p. 924) in [25] and our assumption A1. The second summand converges \mathbb{P} -a.s. to 0 too, which can be seen as follows. From (13) we obtain for n sufficiently large

$$|w_{ni} - 1| \leq \begin{cases} 2 & , \quad i = 1, \dots, \ell_n \\ \frac{\ell_n - 1}{n - \ell_n + 1} & , \quad i = \ell_n + 1, \dots, n - \ell_n \\ 2 & , \quad i = n - \ell_n + 1, \dots, n \end{cases}$$

so that for n sufficiently large

$$\begin{aligned} &\left| \frac{1}{n} \sum_{j=1}^n (w_{ni} - 1) \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) \right| \\ &\leq \frac{\ell_n - 1}{n - \ell_n + 1} \frac{1}{n} \sum_{j=\ell_n+1}^{n-\ell_n} \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) \\ &\quad + 2 \frac{2\ell_n}{n} \frac{1}{2\ell_n} \left(\sum_{j=1}^{\ell_n} \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) + \sum_{j=n-\ell_n+1}^n \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j) \right) \\ &=: S_{2,1}(n) + S_{2,2}(n). \end{aligned}$$

We have seen above that $\frac{1}{n} \sum_{j=1}^n \phi(X_j) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_j)$ converges \mathbb{P} -a.s. to the constant $\mathbb{E}[\phi(X_1) \mathbb{1}_{(y_{i-1}^\varepsilon, y_i^\varepsilon]}(X_1)]$. Since ℓ_n converges to ∞ at a slower rate than n (by assumption A3.), it follows that $S_{2,1}(n)$ converges \mathbb{P} -a.s. to 0. Using the same arguments we obtain that $S_{2,2}(n)$ converges \mathbb{P} -a.s. to 0. Hence $S_2(n)$ converges \mathbb{P} -a.s. to 0. Analogously one can show that $S_1(n)$ converges \mathbb{P} -a.s. to 0. \square

4. Examples for uniformly quasi-Hadamard differentiable functionals

4.1. Average Value at Risk functional

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space and $L^1 = L^1(\Omega, \mathcal{F}, \mathbb{P})$ be the usual L^1 -space. The Average Value at Risk at level $\alpha \in (0, 1)$ is the map $\text{AV@R}_\alpha : L^1 \rightarrow \mathbb{R}$ defined by

$$\text{AV@R}_\alpha(X) := \int_\alpha^1 F_X^\leftarrow(s) ds = - \int_{-\infty}^0 g_\alpha(F_X(x)) dx + \int_0^\infty (1 - g_\alpha(F_X(x))) dx, \quad (21)$$

where $g_\alpha(t) := \frac{1}{1-\alpha} \max\{t - \alpha; 0\}$ and $F_X^\leftarrow(s) := \inf\{x \in \mathbb{R} : F_X(x) \geq s\}$ denotes the left-continuous inverse of the distribution function F_X of X . Note that $\text{AV@R}_\alpha(X) = \mathbb{E}[X | X \geq F_X^\leftarrow(\alpha)]$ when F_X is continuous at $F_X^\leftarrow(\alpha)$, and that AV@R_α is one of the most popular risk measures in practice. In view of the second identity in (21) we may associate with AV@R_α the statistical functional $\mathcal{R}_\alpha : \mathbf{F}_1 \rightarrow \mathbb{R}$ defined by

$$\mathcal{R}_\alpha(F) := - \int_{-\infty}^0 g_\alpha(F(x)) dx + \int_0^\infty (1 - g_\alpha(F(x))) dx, \quad (22)$$

where \mathbf{F}_1 is the set of the distribution functions F_X of all $X \in L^1$. Using the notation introduced at the beginning of Section 3, we obtain the following result.

Proposition 4.1 *Let $F \in \mathbf{F}_1$ and assume that F takes the value $1 - \alpha$ only once. Let \mathcal{S} be the set of all sequences $(G_n) \subseteq \mathbf{F}_1$ with $G_n \rightarrow F$ pointwise. Moreover assume that $\int 1/\phi(x) dx < \infty$. Then the map $\mathcal{R}_\alpha : \mathbf{F}_1 (\subseteq \mathbf{D}) \rightarrow \mathbb{R}$ is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{D}_\phi \langle \mathbf{D}_\phi \rangle$, and the uniform quasi-Hadamard derivative $\dot{\mathcal{R}}_{\alpha;F} : \mathbf{D}_\phi \rightarrow \mathbb{R}$ is given by*

$$\dot{\mathcal{R}}_{\alpha;F}(v) := - \int g'_\alpha(F(x))v(x) dx \quad (23)$$

with $g'_\alpha(t) := \frac{1}{1-\alpha} \mathbb{1}_{(1-\alpha, 1]}(t)$.

Proposition 4.1 shows in particular that for any $F \in \mathbf{F}_1$ which takes the value $1 - \alpha$ only once, the map $\mathcal{R}_\alpha : \mathbf{F}_1 (\subseteq \mathbf{D}) \rightarrow \mathbb{R}$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_\phi \langle \mathbf{D}_\phi \rangle$ (in the sense of part (ii) of Definition 2.1) with uniform quasi-Hadamard derivative given by (23).

Proof (of Proposition 4.1) First of all note that the map $\dot{\mathcal{R}}_{\alpha;F}$ defined in (23) is continuous w.r.t. $\|\cdot\|_\phi$, because

$$|\dot{\mathcal{R}}_{\alpha;F}(v_1) - \dot{\mathcal{R}}_{\alpha;F}(v_2)| \leq \int \frac{1}{1-\alpha} |v_1(x) - v_2(x)| dx \leq \left(\frac{1}{1-\alpha} \int 1/\phi(x) dx \right) \|v_1 - v_2\|_\phi$$

holds for every $v_1, v_2 \in \mathbf{D}_\phi$.

Now, let $((F_n), v, (v_n), (\varepsilon_n))$ be a quadruple with $(F_n) \subseteq \mathbf{F}_1$ satisfying $F_n \rightarrow F$ pointwise, $v \in \mathbf{D}_\phi$, $(v_n) \subseteq \mathbf{D}_\phi$ satisfying $\|v_n - v\|_\phi \rightarrow 0$ and $(F_n + \varepsilon_n v_n) \subseteq \mathbf{F}_1$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \rightarrow 0$. It remains to show that

$$\lim_{n \rightarrow \infty} \left| \frac{\mathcal{R}_\alpha(F_n + \varepsilon_n v_n) - \mathcal{R}_\alpha(F_n)}{\varepsilon_n} - \dot{\mathcal{R}}_{\alpha; F}(v) \right| = 0,$$

that is, in other words, that

$$\lim_{n \rightarrow \infty} \left| \int \left(\frac{g_\alpha(F_n(x)) - g_\alpha((F_n + \varepsilon_n v_n)(x))}{\varepsilon_n} - (-g'_\alpha(F(x))v(x)) \right) dx \right| = 0. \quad (24)$$

Let us denote the integrand of the integral in (24) by $I_n(x)$. In virtue of $F_n \rightarrow F$ pointwise, $\|v_n - v\|_\phi \rightarrow 0$, $\varepsilon_n \rightarrow 0$, and

$$|(F_n + \varepsilon_n v_n)(x) - F(x)| \leq |F_n(x) - F(x)| + \varepsilon_n |v_n(x) - v(x)| + \varepsilon_n |v(x)|,$$

we have $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ and $\lim_{n \rightarrow \infty} (F_n(x) + \varepsilon_n v_n(x)) = F(x)$ for every $x \in \mathbb{R}$. Thus, for every $x \in \mathbb{R}$ with $F(x) < 1 - \alpha$ we obtain $g'_\alpha(F(x))v(x) = 0$ and

$$\frac{g_\alpha(F_n(x)) - g_\alpha((F_n + \varepsilon_n v_n)(x))}{\varepsilon_n} = 0 \quad \text{for sufficiently large } n,$$

i.e. $\lim_{n \rightarrow \infty} I_n(x) = 0$. Moreover for every $x \in \mathbb{R}$ with $F(x) > 1 - \alpha$ we obtain $g'_\alpha(F(x))v(x) = \frac{1}{1-\alpha}v(x)$ and

$$\frac{g_\alpha(F_n(x)) - g_\alpha((F_n + \varepsilon_n v_n)(x))}{\varepsilon_n} = -\frac{v_n(x)}{1-\alpha} \quad \text{for sufficiently large } n,$$

i.e. $\lim_{n \rightarrow \infty} I_n(x) = 0$. Since we assumed that F takes the value $1 - \alpha$ only once, we can conclude that $\lim_{n \rightarrow \infty} I_n(x) = 0$ for Lebesgue-a.e. $x \in \mathbb{R}$. Moreover, by the Lipschitz continuity of g_α with Lipschitz constant $\frac{1}{1-\alpha}$ we have

$$\begin{aligned} |I_n(x)| &= |I_n(x)| \phi(x) \phi(x)^{-1} \\ &= \left| \frac{g_\alpha(F_n(x)) - g_\alpha((F_n + \varepsilon_n v_n)(x))}{\varepsilon_n} + g'_\alpha(F(x))v(x) \right| \phi(x) \phi(x)^{-1} \\ &\leq \frac{1}{1-\alpha} (\|v_n\|_\phi + \|v\|_\phi) \phi(x)^{-1} \\ &\leq \frac{1}{1-\alpha} \left(\sup_{n \in \mathbb{N}} \|v_n\|_\phi + \|v\|_\phi \right) \phi(x)^{-1}. \end{aligned}$$

Since $\sup_{n \in \mathbb{N}} \|v_n\|_\phi < \infty$ (recall $\|v_n - v\|_\phi \rightarrow 0$), the assumption $\int 1/\phi(x) dx < \infty$ ensures that the latter expression provides a Borel measurable majorant of I_n . Now, the Dominated Convergence theorem implies (24). \square

As an immediate consequence of Theorem 3.2, Examples 3.3 and 3.4, and Proposition 4.1 we obtain the following corollary.

Corollary 4.2 *Let $F, \widehat{F}_n, \widehat{F}_n^*, \widehat{C}_n$, and B_F be as in Example 3.3 (S1. or S2.) or as in Example 3.4 respectively, and assume that the assumptions discussed in Example 3.3 or in Example 3.4 respectively are fulfilled for some weight function ϕ with $\int 1/\phi(x) dx < \infty$ (in particular $F \in \mathbf{F}_1$). Then*

$$\sqrt{n}(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F)) \rightsquigarrow \dot{\mathcal{R}}_{\alpha;F}(B_F) \quad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

and

$$\sqrt{n}(\mathcal{R}_\alpha(\widehat{F}_n^*(\omega, \cdot)) - \mathcal{R}_\alpha(\widehat{C}_n(\omega))) \rightsquigarrow \dot{\mathcal{R}}_{\alpha;F}(B_F) \quad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R})), \quad \mathbb{P}\text{-a.e. } \omega.$$

For the bootstrap scheme S1. in Example 3.3 the result of the preceding corollary can be also deduced from Theorem 7 in [16]. According to [17], condition (1) of this theorem is satisfied if there are $0 = a_0 < a_1 < \dots < a_k = 1$ for some $k \in \mathbb{N}$ such that J is Hölder continuous on each interval (a_{i-1}, a_i) , $1 \leq i \leq k$, and the measure dF^{-1} has no mass at the points a_1, \dots, a_{k-1} .

4.2. Compound distribution functional

Let $p = (p_k)_{k \in \mathbb{N}_0}$ be a sequence in \mathbb{R}_+ with $\sum_{k=0}^{\infty} p_k = 1$, so that p specifies the distribution of a count variable N . Let \mathbf{F} denote the set of all distribution functions on \mathbb{R} , and consider the functional $\mathcal{C}_p : \mathbf{F} \rightarrow \mathbf{F}$ defined by

$$\mathcal{C}_p(F) := \sum_{k=0}^{\infty} p_k F^{*k}, \quad (25)$$

where F^{*k} refers to the k -fold convolution of F , that is, $F^{*0} := \mathbb{1}_{[0, \infty)}$ and

$$\begin{aligned} F^{*k}(x) &:= \int F(x - x_{k-1}) dF^{*(k-1)}(x_{k-1}) \\ &= \int \dots \int F(x - x_{k-1} - \dots - x_1) dF(x_1) \dots dF(x_{k-1}) \end{aligned}$$

for $k \in \mathbb{N}$. If $p_m = 1$ for some $m \in \mathbb{N}_0$, then $\mathcal{C}_p(F) = F^{*m}$.

For any $\lambda \geq 0$, let the function $\phi_\lambda : \mathbb{R} \rightarrow [1, \infty)$ be defined by $\phi_\lambda(x) := (1 + |x|)^\lambda$ and denote by $\mathbf{F}_{\phi_\lambda}$ the set of all distribution functions F that satisfy $\int \phi_\lambda(x) dF(x) < \infty$. Using the notation introduced at the beginning of Section 3 and the terminology of part (ii) of Definition 2.1, we obtain the following Proposition 4.3. In the proposition the functional \mathcal{C}_p is restricted to the domain $\mathbf{F}_{\phi_\lambda}$ in order to obtain $\mathbf{D}_{\phi_\lambda}$ as the corresponding trace. The latter will be important for Corollary 4.6.

Proposition 4.3 *Let $\lambda > \lambda' \geq 0$ and $F \in \mathbf{F}_{\phi_\lambda}$. Assume that $\sum_{k=1}^{\infty} p_k k^{(1+\lambda)\vee 2} < \infty$. Then the map $\mathcal{C}_p : \mathbf{F}_{\phi_\lambda}(\subseteq \mathbf{D}) \rightarrow \mathbf{F}(\subseteq \mathbf{D})$ is uniformly quasi-Hadamard differentiable*

at F tangentially to $\mathbf{D}_{\phi_\lambda} \langle \mathbf{D}_{\phi_\lambda} \rangle$ with trace $\mathbf{D}_{\phi_{\lambda'}}$. Moreover, the uniform quasi-Hadamard derivative $\dot{\mathcal{C}}_{p;F} : \mathbf{D}_{\phi_\lambda} \rightarrow \mathbf{D}_{\phi_{\lambda'}}$ is given by

$$\dot{\mathcal{C}}_{p;F}(v)(\cdot) := v * H_{p,F}(\cdot) := \int v(\cdot - x) dH_{p,F}(x), \quad (26)$$

where $H_{p,F} := \sum_{k=1}^{\infty} k p_k F^{*(k-1)}$. In particular, if $p_m = 1$ for some $m \in \mathbb{N}$, then

$$\dot{\mathcal{C}}_{p;F}(v)(\cdot) = m \int v(\cdot - x) dF^{*(m-1)}(x).$$

Proposition 4.3 extends Proposition 4.1 of [23]. Before we prove the proposition, we note that the proposition together with Theorem 3.2 and Examples 3.3 and 3.4 yields the following corollary.

Corollary 4.4 *Let F , \widehat{F}_n , \widehat{F}_n^* , \widehat{C}_n , and B_F be as in Example 3.3 (S1. or S2.) or as in Example 3.4 respectively, and assume that the assumptions discussed in Example 3.3 or in Example 3.4 respectively are fulfilled for some weight function ϕ with $\int 1/\phi(x) dx < \infty$ (in particular $F \in \mathbf{F}_1$). Then for $\lambda' \in (0, \lambda)$*

$$\sqrt{n}(\mathcal{C}_p(\widehat{F}_n) - \mathcal{C}_p(F)) \rightsquigarrow^\circ \dot{\mathcal{C}}_{p;F}(B_F) \quad \text{in } (\mathbf{D}_{\phi_{\lambda'}}, \mathcal{D}_{\phi_{\lambda'}}, \|\cdot\|_{\phi_{\lambda'}})$$

and

$$\sqrt{n}(\mathcal{C}_p(\widehat{F}_n^*(\omega, \cdot)) - \mathcal{C}_p(\widehat{C}_n(\omega))) \rightsquigarrow^\circ \dot{\mathcal{C}}_{p;F}(B_F) \quad \text{in } (\mathbf{D}_{\phi_{\lambda'}}, \mathcal{D}_{\phi_{\lambda'}}, \|\cdot\|_{\phi_{\lambda'}}), \quad \mathbb{P}\text{-a.e. } \omega.$$

To ease the exposition of the proof of Proposition 4.3 we first state a lemma that follows from results given in [23]. In the sequel we will use $f * H$ to denote the function defined by $f * H(\cdot) := \int v(\cdot - x) dH(x)$ for any measurable function f and any distribution function H of a finite (not necessarily probability) Borel measure on \mathbb{R} for which $f * H(\cdot)$ is well defined on \mathbb{R} .

Lemma 4.5 *Let $\lambda > \lambda' \geq 0$, and $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$ and $(G_n) \subseteq \mathbf{F}_{\phi_\lambda}$ be any sequences such that $\|F_n - F\|_{\phi_\lambda} \rightarrow 0$ and $\|G_n - G\|_{\phi_\lambda} \rightarrow 0$ for some $F, G \in \mathbf{F}_{\phi_\lambda}$. Then the following two assertions hold.*

(i) *There exists a constant $C_1 > 0$ such that for every $k, n \in \mathbb{N}$*

$$\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} \leq (2^{\lambda'-1} \vee 1)(1 + k^{\lambda' \vee 1} C_1).$$

(ii) *For every $v \in \mathbf{D}_{\phi_{\lambda'}}$ there exists a constant $C_2 > 0$ such that for every $k, \ell, n \in \mathbb{N}$*

$$\|v * (F_n^{*k} * G_n^{*\ell})\|_{\phi_{\lambda'}} \leq 2^{\lambda'} (1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1)(2 + (k + \ell)^{\lambda' \vee 1} C_2)) \|v\|_{\phi_{\lambda'}}.$$

Proof (i): From (2.4) in [23] we have

$$\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} \leq (2^{\lambda'-1} \vee 1) \left(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} dF_n(x)\right),$$

so that it remains to show that $\int |x|^{\lambda'} dF_n(x)$ is bounded above uniformly in $n \in \mathbb{N}$. The functions $\mathbb{1}_{[0,\infty)} - F_n$ and $\mathbb{1}_{[0,\infty)} - F$ lie in $\mathbf{D}_{\phi_\lambda}$, because $F_n, F \in \mathbf{F}_{\phi_\lambda}$. Along with $\|F_n - F\|_{\phi_\lambda} \rightarrow 0$ this implies $\int |x|^{\lambda'} dF_n(x) \rightarrow \int |x|^{\lambda'} dF(x)$; see Lemma 2.1 in [23]. Therefore, $\int |x|^{\lambda'} dF_n(x) \leq C_1$ for some suitable finite constant $C_1 > 0$ and all $n \in \mathbb{N}$.

(ii): With the help of Lemma 2.3 of [23] (along with $\|F_n^{*k} * G_n^{*\ell}\|_\infty = 1$), Lemma 2.4 of [23], and Equation (2.4) in [23] we obtain

$$\begin{aligned} & \|v * (F_n^{*k} * G_n^{*\ell})\|_{\phi_{\lambda'}} \\ & \leq 2^{\lambda'} \|v\|_{\phi_{\lambda'}} (1 + \|\mathbb{1}_{[0,\infty)} - F_n^{*k} * G_n^{*\ell}\|_{\phi_{\lambda'}}) \\ & \leq 2^{\lambda'} \|v\|_{\phi_{\lambda'}} (1 + 2^{\lambda'} (\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} + \|\mathbb{1}_{(0,\infty)} - G_n^{*\ell}\|_{\phi_{\lambda'}})) \\ & \leq 2^{\lambda'} \|v\|_{\phi_{\lambda'}} \left(1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) \left(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} dF_n(x) + 1 + \ell^{\lambda' \vee 1} \int |x|^{\lambda'} dG_n(x)\right)\right). \end{aligned}$$

So it remains to show that $\int |x|^{\lambda'} dF_n(x)$ and $\int |x|^{\lambda'} dG_n(x)$ are bounded above uniformly in $n \in \mathbb{N}$. But this was already done in the proof of part (i). \square

Proof of Proposition 4.3 First, note that for $G_1, G_2 \in \mathbf{F}_{\phi_\lambda}$ we have

$$\begin{aligned} \|\mathcal{C}_p(G_1) - \mathcal{C}_p(G_2)\|_{\phi_{\lambda'}} & \leq \|\mathcal{C}_p(G_1) - \mathbb{1}_{[0,\infty)}\|_{\phi_{\lambda'}} + \|\mathbb{1}_{[0,\infty)} - \mathcal{C}_p(G_2)\|_{\phi_{\lambda'}} \\ & \leq \int (1 + |x|)^{\lambda'} d\mathcal{C}_p(G_1)(x) + \int (1 + |x|)^{\lambda'} d\mathcal{C}_p(G_2)(x) \end{aligned}$$

by Equation (2.1) in [23]. Moreover, according to Lemma 2.2 in [23] we have that the integrals $\int |x|^{\lambda'} d\mathcal{C}_p(F)(x)$ and $\int |x|^{\lambda'} d\mathcal{C}_p(G)(x)$ are finite under the assumptions of the proposition. Hence, $\mathbf{D}_{\phi_{\lambda'}}$ can indeed be seen as the trace.

Second, we show $(\|\cdot\|_{\phi_\lambda}, \|\cdot\|_{\phi_{\lambda'}})$ -continuity of the map $\dot{\mathcal{C}}_{p;F} : \mathbf{D}_{\phi_\lambda} \rightarrow \mathbf{D}_{\phi_{\lambda'}}$. To this end let $v \in \mathbf{D}_{\phi_\lambda}$ and $(v_n) \subseteq \mathbf{D}_{\phi_\lambda}$ such that $\|v_n - v\|_{\phi_\lambda} \rightarrow 0$. For every $k \in \mathbb{N}$ we have

$$\begin{aligned} & \|p_k k(v_n - v) * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ & \leq 2^{\lambda'} \|v_n - v\|_{\phi_{\lambda'}} p_k k(\|\mathbb{1}_{[0,\infty)}\|_\infty \|F^{*(k-1)}\|_\infty - \|F^{*(k-1)}\|_{\phi_{\lambda'}} + \|F^{*(k-1)}\|_\infty) \\ & = 2^{\lambda'} \|v_n - v\|_{\phi_{\lambda'}} p_k k(\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)}\|_{\phi_{\lambda'}} + 1) \\ & \leq 2^{\lambda'} \|v_n - v\|_{\phi_{\lambda'}} p_k k \left((2^{\lambda'-1} \vee 1) \left(1 + (k-1)^{\lambda' \vee 1} \int |x|^{\lambda'} dF(x)\right) + 1 \right), \end{aligned}$$

where the first and the second inequality follow from Lemma 2.3 and Equation (2.4) in [23] respectively. Hence,

$$\|\dot{\mathcal{C}}_{p;F}(v_n) - \dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} = \|v_n * H_{p,F} - v * H_{p,F}\|_{\phi_{\lambda'}}$$

$$\leq 2^{\lambda'} \|v_n - v\|_{\phi_{\lambda'}} \sum_{k=1}^{\infty} p_k k \left((2^{\lambda'-1} \vee 1) \left(1 + (k-1)^{\lambda' \vee 1} \int |x|^{\lambda'} dF(x) \right) + 1 \right).$$

Now, the series converges due to the assumptions, and $\|v_n - v\|_{\phi_{\lambda}} \rightarrow 0$ implies $\|v_n - v\|_{\phi_{\lambda'}} \rightarrow 0$. Thus $\|\dot{\mathcal{C}}_{p;F}(v_n) - \dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} \rightarrow 0$, which proves continuity.

Third, let $((F_n), v, (v_n), (\varepsilon_n))$ be a quadruple with $(F_n) \subseteq \mathbf{F}_{\phi_{\lambda}}$ satisfying $\|F_n - F\|_{\phi_{\lambda}} \rightarrow 0$, $v \in \mathbf{D}_{\phi_{\lambda}}$, $(v_n) \subseteq \mathbf{D}_{\phi_{\lambda}}$ satisfying $\|v_n - v\|_{\phi_{\lambda}} \rightarrow 0$ and $(F_n + \varepsilon_n v_n) \subseteq \mathbf{F}_{\phi_{\lambda}}$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \rightarrow 0$. It remains to show that

$$\lim_{n \rightarrow \infty} \left\| \frac{\mathcal{C}_p(F_n + \varepsilon_n v_n) - \mathcal{C}_p(F_n)}{\varepsilon_n} - \dot{\mathcal{C}}_{p;F}(v) \right\|_{\phi_{\lambda'}} = 0.$$

To do so, define for $k \in \mathbb{N}_0$ a map $H_k : \mathbf{F} \times \mathbf{F} \rightarrow \mathbf{F}$ by

$$H_k(G_1, G_2) := \sum_{j=0}^{k-1} G_1^{*(k-1-j)} * G_2^{*j}.$$

with the usual convention that the sum over the empty sum equals zero. We find that for every $M \in \mathbb{N}$

$$\begin{aligned} & \left\| \frac{\mathcal{C}_p(F_n + \varepsilon_n v_n) - \mathcal{C}_p(F_n)}{\varepsilon_n} - \dot{\mathcal{C}}_{p;F}(v) \right\|_{\phi_{\lambda'}} \\ &= \left\| \frac{1}{\varepsilon_n} \left(\sum_{k=0}^{\infty} p_k (F_n + \varepsilon_n v_n)^{*k} - \sum_{k=0}^{\infty} p_k F_n^{*k} \right) - \dot{\mathcal{C}}_{p;F}(v) \right\|_{\phi_{\lambda'}} \\ &= \left\| \frac{1}{\varepsilon_n} \left(\sum_{k=1}^{\infty} (p_k (F_n + \varepsilon_n v_n)^{*k} - p_k F_n^{*k}) \right) - \dot{\mathcal{C}}_{p;F}(v) \right\|_{\phi_{\lambda'}} \\ &= \left\| \sum_{k=1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) - \dot{\mathcal{C}}_{p;F}(v) \right\|_{\phi_{\lambda'}} \\ &\leq \left\| \sum_{k=M+1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) \right\|_{\phi_{\lambda'}} + \left\| \sum_{k=1}^M p_k (v_n - v) * H_k(F_n + \varepsilon_n v_n, F_n) \right\|_{\phi_{\lambda'}} \\ &\quad + \left\| v * \sum_{k=M+1}^{\infty} k p_k F_n^{*(k-1)} \right\|_{\phi_{\lambda'}} + \left\| \sum_{k=1}^M p_k v * H_k(F_n + \varepsilon_n v_n, F_n) - k p_k v * F_n^{*(k-1)} \right\|_{\phi_{\lambda'}} \\ &=: S_1(n, M) + S_2(n, M) + S_3(M) + S_4(n, M), \end{aligned}$$

where for the third “=” we use the fact that for $G_1, G_2 \in \mathbf{F}$

$$(G_1 - G_2) * H_k(G_1, G_2) = G_1^{*k} - G_2^{*k}. \quad (27)$$

By part (ii) of Lemma 4.5 (this lemma can be applied since $\|F_n + \varepsilon_n v_n - F\|_{\phi_{\lambda}} \rightarrow 0$) there exists a constant $C_2 > 0$ such that for all $n \in \mathbb{N}$

$$S_1(n, M) = \left\| \sum_{k=M+1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) \right\|_{\phi_{\lambda'}}$$

$$\leq 2^{\lambda'} \|v_n\|_{\phi_{\lambda'}} \sum_{k=M+1}^{\infty} p_k k (1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) (2 + (k-1)^{\lambda' \vee 1} C_2)). \quad (28)$$

Since $\lambda' < \lambda$ and $\|v_n - v\|_{\phi_{\lambda}} \rightarrow 0$, we have $\|v_n\|_{\phi_{\lambda'}} \leq K_1$ for some finite constant $K_1 > 0$ and all $n \in \mathbb{N}$. Hence, the right-hand side of (28) can be made arbitrarily small by choosing M large enough. That is, $S_1(n, M)$ can be made arbitrarily small uniformly in $n \in \mathbb{N}$ by choosing M large enough.

Furthermore, it is demonstrated in the proof of Proposition 4.1 of [23] that $S_3(M)$ can be made arbitrarily small by choosing M large enough.

Next, applying again part (ii) of Lemma 4.5 we obtain

$$\begin{aligned} S_2(n, M) &= \left\| \sum_{k=1}^M p_k (v_n - v) * H_k(F_n + \varepsilon_n v_n, F_n) \right\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \sum_{k=1}^M p_k k \|v_n - v\|_{\phi_{\lambda'}} (1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) (2 + (k-1)^{\lambda' \vee 1} C_2)). \end{aligned}$$

Using $\|v_n - v\|_{\phi_{\lambda'}} \leq \|v_n - v\|_{\phi_{\lambda}} \rightarrow 0$ this term tends to zero as $n \rightarrow \infty$ for a given M .

It remains to consider the summand

$$\begin{aligned} S_4(n, M) &= \left\| \sum_{k=1}^M p_k v * H_k(F_n + \varepsilon_n v_n, F_n) - k p_k v * F^{*(k-1)} \right\|_{\phi_{\lambda'}} \\ &= \left\| \sum_{k=1}^M p_k \sum_{\ell=0}^{k-1} \left(v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)} \right) \right\|_{\phi_{\lambda'}}. \end{aligned}$$

We will show that for M fixed this term can be made arbitrarily small by letting $n \rightarrow \infty$. This would follow if for every given $k \in \{1, \dots, M\}$ and $\ell \in \{0, \dots, k-1\}$ the expression

$$\|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)}\|_{\phi_{\lambda'}}$$

could be made arbitrarily small by letting $n \rightarrow \infty$. For every such k and ℓ we can find a linear combination of indicator functions of the form $\mathbb{1}_{[a,b)}$, $-\infty < a < b < \infty$, which we denote by \tilde{v} , such that

$$\begin{aligned} &\|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\leq \|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - \tilde{v} * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} \\ &\quad + \|\tilde{v} * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - \tilde{v} * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\quad + \|\tilde{v} * F^{*(k-1)} - v * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ &\quad + c(\lambda', \tilde{v}) \|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\quad + 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|F^{*(k-1)}\|_{\phi_{\lambda'}} + 1) \end{aligned} \quad (29)$$

for some suitable finite constant $c(\lambda', \tilde{v}) > 0$ depending only on λ' and \tilde{v} . The first inequality in (29) is obvious (and holds for any $\tilde{v} \in \mathbf{D}_{\phi_{\lambda'}}$). The second inequality in (29) is obtained by applying Lemma 2.3 of [23] to the first summand (noting that $\|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_\infty = 1$; recall $F_n + \varepsilon_n v_n \in \mathbf{F}$), by applying Lemma 4.3 of [23] to the second summand (which requires that \tilde{v} is as described above), and by applying Lemma 2.3 of [23] to the third summand.

We now consider the three summands on the right-hand side of (29) separately. We start with the third term. Since $v \in \mathbf{D}_{\phi_\lambda}$, Lemma 4.2 of [23] ensures that we may assume that \tilde{v} is chosen such that $\|\tilde{v} - v\|_{\phi_{\lambda'}}$ is arbitrarily small. Hence, for fixed M the third summand in (29) can be made arbitrarily small.

We next consider the the second summand in (29). Obviously,

$$\begin{aligned} & \|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F_n^{*(k-1)}\|_{\phi_{\lambda'}} \\ &= \|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F_n^{*(k-1)} + F_n^{*(k-1)} - F_n^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\leq \|((F_n + \varepsilon_n v_n)^{*(k-1-\ell)} - F_n^{*(k-1-\ell)}) * F_n^{*\ell}\|_{\phi_{\lambda'}} + \|F_n^{*(k-1)} - F_n^{*(k-1)}\|_{\phi_{\lambda'}}. \end{aligned} \quad (30)$$

We start by considering the first summand in (30). In view of (27) it can be written as

$$\begin{aligned} & \|((F_n + \varepsilon_n v_n)^{*(k-1-\ell)} - F_n^{*(k-1-\ell)}) * F_n^{*\ell}\|_{\phi_{\lambda'}} \\ &= \|((F_n + \varepsilon_n v_n - F_n) * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n)) * F_n^{*\ell}\|_{\phi_{\lambda'}} \\ &= \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n)) * F_n^{*\ell}\|_{\phi_{\lambda'}}. \end{aligned}$$

Applying Lemma 2.3 of [23] with $f = \varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n)$ and $H = F_n^{*\ell}$ we obtain

$$\begin{aligned} & \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n)) * F_n^{*\ell}\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)}\|_{F_n^{*\ell}} - F_n^{*\ell}\|_{\phi_{\lambda'}} + \|F_n^{*\ell}\|_\infty) \\ &= 2^{\lambda'} \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ &\leq 2^{\lambda'} \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} \{(2^{\lambda'-1} \vee 1)(1 + \ell^{\lambda' \vee 1} C_1) + 1\}, \end{aligned} \quad (31)$$

where we applied part (i) of Lemma 4.5 to $\|\mathbb{1}_{[0,\infty)} - F_n^{*\ell}\|_{\phi_{\lambda'}}$ to obtain the last inequality. Hence for the left-hand side of (31) to go to zero as $n \rightarrow \infty$ it suffices to show that $\|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} \rightarrow 0$ as $n \rightarrow \infty$. The latter follows from

$$\begin{aligned} & \|(\varepsilon_n v_n * H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} (k - \ell - 1) \varepsilon_n \|v_n\|_{\phi_{\lambda'}} (1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) (2 + ((k - \ell - 2))^{\lambda' \vee 1} C_2)), \end{aligned} \quad (32)$$

where we applied part (ii) of Lemma 4.5 with $v = \varepsilon_n v_n$ to all summands in $H_{k-1-\ell}(F_n + \varepsilon_n v_n, F_n)$. For every k and $\ell \in \{0, \dots, k-1\}$ this expression goes indeed to zero as $n \rightarrow \infty$, because, as mentioned before, $\|v_n\|_{\phi_{\lambda'}}$ is uniformly bounded in $n \in \mathbb{N}$, and we

have $\varepsilon_n \rightarrow 0$. Next we consider the second summand in (30). Applying (27) to $F_n^{*(k-1)}$ and $F^{*(k-1)}$ and subsequently part (ii) of Lemma 4.5 to the summands in $H_{k-1}(F_n, F)$ we have

$$\|F_n^{*(k-1)} - F^{*(k-1)}\|_{\phi_{\lambda'}} \leq 2^{\lambda'}(k-1)\|F_n - F\|_{\phi_{\lambda'}}(1 + 2^{\lambda'}(2^{\lambda'-1} \vee 1)(2 + ((k-2))^{\lambda' \vee 1} C_2)).$$

Clearly for every k this term goes to zero 0 as $n \rightarrow \infty$, because $\|F_n - F\|_{\phi_{\lambda'}} \leq \|F_n - F\|_{\phi_{\lambda}} \rightarrow 0$ as $n \rightarrow \infty$ by assumption. This together with the fact that (31) goes to zero 0 as $n \rightarrow \infty$ shows that (30) goes to zero in $\|\cdot\|_{\phi_{\lambda'}}$ as $n \rightarrow \infty$. Therefore, the second summand in (29) goes to zero as $n \rightarrow \infty$.

It remains to consider the first term in (29). We find

$$\begin{aligned} & 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ & \leq 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ & \leq 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)} + F^{*(k-1)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ & \leq 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)}\|_{\phi_{\lambda'}} + \|F^{*(k-1)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1) \\ & \leq 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (2^{\lambda'-1} \vee 1) \left(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} dF(x)\right) \\ & \quad + 2^{\lambda'} \|\tilde{v} - v\|_{\phi_{\lambda'}} (\|F^{*(k-1)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1), \end{aligned} \tag{33}$$

where for the last inequality we used Formula (2.4) of [23]. In the lines following (30) we showed that $\|F^{*(k-1)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}}$ goes to zero as $n \rightarrow \infty$ for every k and $\ell \in \{0, \dots, k-1\}$. Hence for every such k and ℓ , it is uniformly bounded in $n \in \mathbb{N}$. Therefore we can make (33) arbitrarily small by making $\|\tilde{v} - v\|_{\phi_{\lambda'}}$ small which, as mentioned above, is possible according to Lemma 4.2 of [23]. This finishes the proof. \square

4.3. Composition of Average Value at Risk functional and compound distribution functional

Here we consider the composition of the Average Value at Risk functional \mathcal{R}_{α} defined in (22) and the compound distribution functional \mathcal{C}_p defined in (25). As a consequence of Propositions 4.1 and 4.3 we obtain the following Corollary 4.6. Note that, for any $\lambda > 1$, Lemma 2.2 in [23] yields $\mathcal{C}_p(\mathbf{F}_{\phi_{\lambda}}) \subseteq \mathbf{F}_1$ so that the composition $\mathcal{R}_{\alpha} \circ \mathcal{C}_p$ is well defined on $\mathbf{F}_{\phi_{\lambda}}$.

Corollary 4.6 *Assume that $\sum_{k=1}^{\infty} p_k k^{(1+\lambda) \vee 2} < \infty$. Let $\lambda > 1$, $F \in \mathbf{F}_{\phi_{\lambda}}$, and assume that $\mathcal{C}_p(F)$ takes the value $1 - \alpha$ only once. Then the map $T_{\alpha,p} := \mathcal{R}_{\alpha} \circ \mathcal{C}_p : \mathbf{F}_{\phi_{\lambda}}(\subseteq \mathbf{D}) \rightarrow \mathbb{R}$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_{\phi_{\lambda}} \langle \mathbf{D}_{\phi_{\lambda}} \rangle$, and the*

uniform quasi-Hadamard derivative $\dot{T}_{\alpha,p;F} : \mathbf{D}_{\phi_\lambda} \rightarrow \mathbb{R}$ is given by $\dot{T}_{\alpha,p;F} = \dot{\mathcal{R}}_{\alpha;\mathcal{C}_p(F)} \circ \dot{\mathcal{C}}_{p;F}$, i.e.

$$\dot{T}_{\alpha,p;F}(v) = \int g'_\alpha(\mathcal{C}_p(F)(x))(v * H_{p,F})(x) dx \quad \text{for all } v \in \mathbf{D}_{\phi_\lambda}$$

with g'_α and $v * H_{p,F}$ as in Proposition 4.1 and 4.3, respectively.

Proof We intend to apply Lemma A.5 to $H = \mathcal{C}_p : \mathbf{F}_{\phi_\lambda} \rightarrow \mathbf{F}_1$ and $\tilde{H} = \mathcal{R}_\alpha : \mathbf{F}_1 \rightarrow \mathbb{R}$. To verify that the assumptions of the lemma are fulfilled, we first recall from the comment directly before Corollary 4.6 that $\mathcal{C}_p(\mathbf{F}_{\phi_\lambda}) \subseteq \mathbf{F}_1$. It remains to show that the assumptions (a)–(c) of Lemma A.5 are fulfilled. According to Proposition 4.3 we have that for every $\lambda' \in (1, \lambda)$ the functional \mathcal{C}_p is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_{\phi_\lambda} \langle \mathbf{D}_{\phi_\lambda} \rangle$ with trace $\mathbf{D}_{\phi_{\lambda'}}$, which is the first part of assumption (b). The second part of assumption (b) means $\dot{\mathcal{C}}_{p;F}(\mathbf{D}_{\phi_\lambda}) \subseteq \mathbf{D}_{\phi_{\lambda'}}$ and follows from

$$\begin{aligned} \|\dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} &= \left\| v * \sum_{k=1}^{\infty} p_k k F^{*(k-1)} \right\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \|v\|_{\phi_{\lambda'}} \sum_{k=1}^{\infty} p_k k \left(1 + (2^{\lambda'-1} \vee 1) \left(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} dF(x) \right) \right) \end{aligned}$$

(for which we applied Lemma 2.3 and Inequality (2.4) in [23]), the convergence of the latter series (which holds by assumption), and $\|v\|_{\phi_{\lambda'}} \leq \|v\|_{\phi_\lambda} < \infty$. Further, it follows from Proposition 4.1 that the map \mathcal{R}_α is uniformly quasi-Hadamard differentiable tangentially to $\mathbf{D}_{\phi_{\lambda'}} \langle \mathbf{D}_{\phi_{\lambda'}} \rangle$ at every distribution function of $\mathbf{F}_{\phi_{\lambda'}}$ that takes the value $1 - \alpha$ only once. This is assumption (c) of Lemma A.5.

It remains to show that also assumption (a) of Lemma A.5 holds true. In the present setting assumption (a) means that for every sequence $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$ with $\|F_n - F\|_{\phi_\lambda} \rightarrow 0$ we have $\mathcal{C}_p(F_n) \rightarrow \mathcal{C}_p(F)$ pointwise. We will show that we even have $\|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} \rightarrow 0$. So let $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$. Then

$$\begin{aligned} \|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} &= \left\| \sum_{k=1}^{\infty} p_k (F_n^{*k} - F^{*k}) \right\|_{\phi_{\lambda'}} \\ &= \left\| (F_n - F) * \sum_{k=1}^{\infty} p_k H_k(F_n, F) \right\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \|F_n - F\|_{\phi_{\lambda'}} \sum_{k=1}^{\infty} p_k k \left(1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) (2 + (k-1)^{\lambda' \vee 1} C_2) \right), \end{aligned}$$

where we used (27) for the second “=” and applied part (ii) of Lemma 4.5 to the summands of H_k to obtain the latter inequality. Since the series converges, we obtain $\|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} \rightarrow 0$ when assuming $\|F_n - F\|_{\phi_\lambda} \rightarrow 0$. \square

As an immediate consequence of Theorem 3.2, Examples 3.3 and 3.4, and Corollary 4.6 we obtain the following corollary.

Corollary 4.7 *Let $F, \widehat{F}_n, \widehat{F}_n^*, \widehat{C}_n$, and B_F be as in Example 3.3 (S1. or S2.) or as in Example 3.4 respectively, and assume that the assumptions discussed in Example 3.3 or in Example 3.4 respectively are fulfilled for some weight function ϕ with $\int 1/\phi(x) dx < \infty$ (in particular $F \in \mathbf{F}_1$). Then*

$$\sqrt{n}(T_{\alpha,p}(\widehat{F}_n) - T_{\alpha,p}(F)) \rightsquigarrow \dot{T}_{\alpha,p;F}(B_F) \quad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

and

$$\sqrt{n}(T_{\alpha,p}(\widehat{F}_n^*(\omega, \cdot)) - T_{\alpha,p}(\widehat{C}_n(\omega))) \rightsquigarrow \dot{T}_{\alpha,p;F}(B_F) \quad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R})), \quad \mathbb{P}\text{-a.e. } \omega.$$

A. Convergence in distribution^o

Let (\mathbf{E}, d) be a metric space and \mathcal{B}° be the σ -algebra on \mathbf{E} generated by the open balls $B_r(x) := \{y \in \mathbf{E} : d(x, y) < r\}$, $x \in \mathbf{E}$, $r > 0$. We will refer to \mathcal{B}° as *open-ball σ -algebra*. If (\mathbf{E}, d) is separable, then \mathcal{B}° coincides with the Borel σ -algebra \mathcal{B} . If (\mathbf{E}, d) is not separable, then \mathcal{B}° might be strictly smaller than \mathcal{B} and thus a continuous real-valued function on \mathbf{E} is not necessarily $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$ -measurable. Let C_b° be the set of all bounded, continuous and $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$ -measurable real-valued functions on \mathbf{E} , and \mathcal{M}_1° be the set of all probability measures on $(\mathbf{E}, \mathcal{B}^\circ)$.

Let X_n be an $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable on some probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ for every $n \in \mathbb{N}_0$. Then the sequence $(X_n) = (X_n)_{n \in \mathbb{N}}$ is said to *converge in distribution^o* to X_0 if

$$\int f d\mathbb{P} \circ X_n^{-1} \longrightarrow \int f d\mathbb{P}_0 \circ X_0^{-1} \quad \text{for all } f \in C_b^\circ.$$

In this case, we write $X_n \rightsquigarrow^\circ X_0$. This is the same as saying that the sequence $(\mathbb{P}_n \circ X_n^{-1})$ converges to $\mathbb{P}_0 \circ X_0^{-1}$ in the weak^o topology on \mathcal{M}_1° ; for details see the Appendix A of [7]. It is worth mentioning that two probability measures $\mu, \nu \in \mathcal{M}_1^\circ$ coincide if $\mu[\mathbf{E}_0] = \nu[\mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$ and $\int f d\mu = \int f d\nu$ for all uniformly continuous $f \in C_b^\circ$; see, for instance, [8, Theorem 6.2].

In the Appendices A–C in [7] several properties of convergence in distribution^o (and weak^o convergence) have been discussed. The following two subsections complement this discussion.

A.1. Slutsky-type results for the open-ball σ -algebra

For a sequence (X_n) of $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variables that are all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the sequence (X_n) is said to *converge in probability^o* to X_0 if the mappings $\omega \mapsto d(X_n(\omega), X_0(\omega))$, $n \in \mathbb{N}$, are $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$ -measurable and satisfy

$$\lim_{n \rightarrow \infty} \mathbb{P}[d(X_n, X_0) \geq \varepsilon] = 0 \quad \text{for all } \varepsilon > 0. \quad (34)$$

In this case, we write $X_n \rightarrow^{\mathbf{P}, \circ} X_0$. The superscript \circ points to the fact that measurability of the mapping $\omega \mapsto d(X_n(\omega), X_0(\omega))$ is a requirement of the definition (and not automatically valid). Note however that in the specific situation where $X_0 \equiv x_0$ for some $x_0 \in \mathbf{E}$, measurability of the mapping $\omega \mapsto d(X_n(\omega), X_0(\omega))$ does hold; cf. Lemma B.3 in [7]. Also note that the measurability always holds when (\mathbf{E}, d) is separable; in this case we also write $\rightarrow^{\mathbf{P}}$ instead of $\rightarrow^{\mathbf{P}, \circ}$.

Theorem A.1 *Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and assume that the mapping $\omega \mapsto d(X_n(\omega), Y_n(\omega))$ is $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$ -measurable for every $n \in \mathbb{N}$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Then $X_n \rightsquigarrow^\circ X_0$ and $d(X_n, Y_n) \rightarrow^{\mathbf{P}} 0$ together imply $Y_n \rightsquigarrow^\circ X_0$.*

Proof In view of $X_n \rightsquigarrow^\circ X$, we obtain for every fixed $f \in \text{BL}_1^\circ$

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \left| \int f d\mathbb{P}_{Y_n} - \int f d\mathbb{P}_{X_0} \right| \\ & \leq \limsup_{n \rightarrow \infty} \left| \int f d\mathbb{P}_{Y_n} - \int f d\mathbb{P}_{X_n} \right| + \limsup_{n \rightarrow \infty} \left| \int f d\mathbb{P}_{X_n} - \int f d\mathbb{P}_{X_0} \right| \\ & \leq \limsup_{n \rightarrow \infty} \int |f(Y_n) - f(X_n)| d\mathbb{P}. \end{aligned}$$

Since f lies in BL_1° and we assumed $d(X_n, Y_n) \rightarrow^{\mathbf{P}} 0$, we also have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \int |f(Y_n) - f(X_n)| d\mathbb{P} & \leq \limsup_{n \rightarrow \infty} \left| \int |f(Y_n) - f(X_n)| \mathbb{1}_{\{d(X_n, Y_n) \geq \varepsilon\}} d\mathbb{P} + 2\varepsilon \right| \\ & \leq 2 \limsup_{n \rightarrow \infty} \mathbb{P}[d(X_n, Y_n) \geq \varepsilon] + 2\varepsilon \end{aligned}$$

for every $\varepsilon > 0$. Thus $\limsup_{n \rightarrow \infty} \int |f(Y_n) - f(X_n)| d\mathbb{P} = 0$ which together with the Portmanteau theorem (in the form of [7, Theorem A.4]) implies the claim. \square

Set $\overline{\mathbf{E}} := \mathbf{E} \times \mathbf{E}$ and let $\overline{\mathcal{B}}^\circ$ be the σ -algebra on $\overline{\mathbf{E}}$ generated by the open balls w.r.t. the metric

$$\overline{d}((x_1, x_2), (y_1, y_2)) := \max\{d(x_1, y_1); d(x_2, y_2)\}.$$

Recall that $\overline{\mathcal{B}}^\circ \subseteq \mathcal{B}^\circ \otimes \mathcal{B}^\circ$, where the inclusion may be strict.

Corollary A.2 *Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let $y_0 \in \mathbf{E}_0$. Let $(\tilde{\mathbf{E}}, \tilde{d})$ be a metric space equipped with the corresponding open-ball σ -algebra $\tilde{\mathcal{B}}^\circ$. Then $X_n \rightsquigarrow^\circ X_0$ and $Y_n \rightarrow^{\mathbf{P}, \circ} y_0$ together imply*

$$(i) \quad (X_n, Y_n) \rightsquigarrow^\circ (X_0, y_0).$$

(ii) $h(X_n, Y_n) \rightsquigarrow^\circ h(X_0, y_0)$ for every continuous and $(\overline{\mathcal{B}}, \widetilde{\mathcal{B}}^\circ)$ -measurable $h : \overline{\mathbf{E}} \rightarrow \widetilde{\mathbf{E}}$.

Proof Assertion (ii) is an immediate consequence of assertion (i) and the Continuous Mapping theorem in the form of [8, Theorem 6.4]; take into account that (X_0, y_0) takes values only in $\overline{\mathbf{E}}_0 := \mathbf{E}_0 \times \mathbf{E}_0$ and that $\mathbf{E}_0 \times \mathbf{E}_0$ is separable w.r.t. \overline{d} . Thus it suffices to show assertion (i). First note that we have

$$(X_n, y_0) \rightsquigarrow^\circ (X_0, y_0). \quad (35)$$

Indeed, for every $f \in \overline{\mathcal{C}}_b^\circ$ (with $\overline{\mathcal{C}}_b^\circ$ the set of all bounded, continuous and $(\overline{\mathcal{B}}, \mathcal{B}(\mathbb{R}))$ -measurable real-valued functions on $\overline{\mathbf{E}}$) we have $\lim_{n \rightarrow \infty} \int f(X_n, y_0) d\mathbb{P} = \int f(X_0, y_0) d\mathbb{P}_0$ by the assumption $X_n \rightsquigarrow^\circ X_0$ and the fact that the mapping $x \mapsto f(x, y_0)$ lies in C_b° (the latter was shown in the proof of Theorem 3.1 in [7]).

Second, the distance $\overline{d}((X_n, Y_n), (X_n, y_0)) = d(Y_n, y_0)$ is $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$ -measurable for every $n \in \mathbb{N}$, because Y_n is $(\mathcal{F}, \mathcal{B}^\circ)$ -measurable and $x \mapsto d(x, y_0)$ is $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$ -measurable (due to Lemma B.3 in [7]). Along with $Y_n \xrightarrow{\mathbb{P}, \circ} y_0$ we obtain in particular that $\overline{d}((X_n, Y_n), (X_n, y_0)) \xrightarrow{\mathbb{P}} 0$. Together with (35) and Theorem A.1 (applied to $X'_n := (X_n, y_0)$, $X'_0 := (X_0, y_0)$, $Y'_n := (X_n, Y_n)$) this implies $(X_n, Y_n) \rightsquigarrow^\circ (X_0, y_0)$; take into account again that (X_0, y_0) takes values only in $\overline{\mathbf{E}}_0 := \mathbf{E}_0 \times \mathbf{E}_0$ and that $\mathbf{E}_0 \times \mathbf{E}_0$ is separable w.r.t. \overline{d} . \square

Corollary A.3 *Let $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ be a normed vector space and d be the induced metric defined by $d(x_1, x_2) := \|x_1 - x_2\|_{\mathbf{E}}$. Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let $y_0 \in \mathbf{E}_0$. Assume that the map $h : \overline{\mathbf{E}} \rightarrow \widetilde{\mathbf{E}}$ defined by $h(x_1, x_2) := x_1 + x_2$ is $(\overline{\mathcal{B}}, \mathcal{B}^\circ)$ -measurable. Then $X_n \rightsquigarrow^\circ X_0$ and $Y_n \xrightarrow{\mathbb{P}, \circ} y_0$ together imply $X_n + Y_n \rightsquigarrow^\circ X_0 + y_0$.*

Proof The assertion is an immediate consequence of Corollary A.2 and the fact that h is clearly continuous (w.r.t. \overline{d} and the Euclidean distance $|\cdot|$). \square

A.2. Delta-method and chain rule for uniformly quasi-Hadamard differentiable maps

Now assume that \mathbf{E} is a subspace of a vector space \mathbf{V} . Let $\|\cdot\|_{\mathbf{E}}$ be a norm on \mathbf{E} and assume that the metric d is induced by $\|\cdot\|_{\mathbf{E}}$. Let $\widetilde{\mathbf{V}}$ be another vector space and $\widetilde{\mathbf{E}} \subseteq \widetilde{\mathbf{V}}$ be any subspace. Let $\|\cdot\|_{\widetilde{\mathbf{E}}}$ be a norm on $\widetilde{\mathbf{E}}$ and $\widetilde{\mathcal{B}}^\circ$ be the corresponding open-ball σ -algebra on $\widetilde{\mathbf{E}}$. Let $0_{\widetilde{\mathbf{E}}}$ denote the null in $\widetilde{\mathbf{E}}$. Moreover, let $\widetilde{\widetilde{\mathbf{E}}} := \widetilde{\mathbf{E}} \times \widetilde{\mathbf{E}}$ and $\widetilde{\widetilde{\mathcal{B}}}^\circ$ be the σ -algebra on $\widetilde{\widetilde{\mathbf{E}}}$ generated by the open balls w.r.t. the metric $\widetilde{\widetilde{d}}((\widetilde{x}_1, \widetilde{x}_2), (\widetilde{y}_1, \widetilde{y}_2)) := \max\{\|\widetilde{x}_1 - \widetilde{y}_1\|_{\widetilde{\mathbf{E}}}; \|\widetilde{x}_2 - \widetilde{y}_2\|_{\widetilde{\mathbf{E}}}\}$.

Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ be a probability space and $\widehat{T}_n : \Omega_n \rightarrow \mathbf{V}$ be any map for every $n \in \mathbb{N}$. Recall that \rightsquigarrow° and $\rightarrow^{\mathbf{P}, \circ}$ refer to convergence in distribution^o and convergence in probability^o, respectively. Moreover recall Definition 2.1 of quasi-Hadamard differentiability.

Theorem A.4 *Let $H : \mathbf{V}_H \rightarrow \widetilde{\mathbf{E}}$ be a map defined on some $\mathbf{V}_H \subseteq \mathbf{V}$. Let $\mathbf{E}_0 \in \mathcal{B}^\circ$ be some $\|\cdot\|_{\mathbf{E}}$ -separable subset of \mathbf{E} . Let $(\theta_n) \subseteq \mathbf{V}_H$ and define the singleton set $\mathcal{S} := \{(\theta_n)\}$. Let (a_n) be a sequence of positive real numbers tending to ∞ , and consider the following conditions:*

(a) \widehat{T}_n takes values only in \mathbf{V}_H .

(b) $a_n(\widehat{T}_n - \theta_n)$ takes values only in \mathbf{E} , is $(\mathcal{F}_n, \mathcal{B}^\circ)$ -measurable and satisfies

$$a_n(\widehat{T}_n - \theta_n) \rightsquigarrow^\circ \xi \quad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}) \quad (36)$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$ -valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.

(c) $a_n(H(\widehat{T}_n) - H(\theta_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable.

(d) The map H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0 \langle \mathbf{E} \rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_S : \mathbf{E}_0 \rightarrow \widetilde{\mathbf{E}}$.

(e) $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n) = (\Omega, \mathcal{F}, \mathbb{P})$ for all $n \in \mathbb{N}$.

(f) The uniform quasi-Hadamard derivative \dot{H}_S can be extended to \mathbf{E} such that the extension $\dot{H}_S : \mathbf{E} \rightarrow \widetilde{\mathbf{E}}$ is continuous at every point of \mathbf{E}_0 and $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable.

(g) The map $h : \widetilde{\mathbf{E}} \rightarrow \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\widetilde{\mathcal{B}}^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable.

Then the following two assertions hold:

(i) If conditions (a)–(d) hold true, then $\dot{H}_S(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$ -measurable and

$$a_n(H(\widehat{T}_n) - H(\theta_n)) \rightsquigarrow^\circ \dot{H}_S(\xi) \quad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}).$$

(ii) If conditions (a)–(g) hold true, then

$$a_n(H(\widehat{T}_n) - H(\theta_n)) - \dot{H}_S(a_n(\widehat{T}_n - \theta_n)) \rightarrow^{\mathbf{P}, \circ} 0_{\widetilde{\mathbf{E}}} \quad \text{in } (\widetilde{\mathbf{E}}, \|\cdot\|_{\widetilde{\mathbf{E}}}). \quad (37)$$

Proof The proof is very similar to the proof of Theorem C.4 in [7].

(i): For every $n \in \mathbb{N}$, let $\mathbf{E}_n := \{x_n \in \mathbf{E} : \theta_n + a_n^{-1}x_n \in \mathbf{V}_H\}$ and define the map $h_n : \mathbf{E}_n \rightarrow \widetilde{\mathbf{E}}$ by

$$h_n(x_n) := \frac{H(\theta_n + a_n^{-1}x_n) - H(\theta_n)}{a_n^{-1}}.$$

Moreover, define the map $h_0 : \mathbf{E}_0 \rightarrow \widetilde{\mathbf{E}}$ by

$$h_0(x) := \dot{H}_{\mathcal{S}}(x).$$

Now, the claim would follow by the extended Continuous Mapping theorem in the form of Theorem C.1 in [7] applied to the functions h_n , $n \in \mathbb{N}_0$, and the random variables $\xi_n := a_n(\widehat{T}_n - \theta_n)$, $n \in \mathbb{N}$, and $\xi_0 := \xi$ if we can show that the assumptions of Theorem C.1 in [7] are satisfied. First, by assumption (a) and the last part of assumption (b) we have $\xi_n(\Omega_n) \subseteq \mathbf{E}_n$ and $\xi_0(\Omega_0) \subseteq \mathbf{E}_0$. Second, by assumption (c) we have that $h_n(\xi_n) = a_n(H(\widehat{T}_n) - H(\theta_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable. Third, the map h_0 is continuous by the definition of the quasi-Hadamard derivative. Thus h_0 is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable, because the trace σ -algebra $\mathcal{B}_0^\circ := \mathcal{B}^\circ \cap \mathbf{E}_0$ coincides with the Borel σ -algebra on \mathbf{E}_0 (recall that \mathbf{E}_0 is separable). In particular, $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$ -measurable. Fourth, condition (a) of Theorem C.1 in [7] holds by assumption (b). Fifth, condition (b) of Theorem C.1 in [7] is ensured by assumption (d).

(ii): For every $n \in \mathbb{N}$, let \mathbf{E}_n and h_n be as above and define the map $\bar{h}_n : \mathbf{E}_n \rightarrow \widetilde{\mathbf{E}}$ by

$$\bar{h}_n(x_n) := (h_n(x_n), \dot{H}_{\mathcal{S}}(x_n)).$$

Moreover, define the map $\bar{h}_0 : \mathbf{E}_0 \rightarrow \widetilde{\mathbf{E}}$ by

$$\bar{h}_0(x) := (h_0(x), \dot{H}_{\mathcal{S}}(x)) = (\dot{H}_{\mathcal{S}}(x), \dot{H}_{\mathcal{S}}(x)).$$

We will first show that

$$\bar{h}_n(a_n(X_n - x)) \rightsquigarrow^\circ \bar{h}_0(X_0) \quad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \widetilde{d}). \quad (38)$$

For (38) it suffices to show that the assumption of the extended Continuous Mapping theorem in the form of Theorem C.1 in [7] applied to the functions \bar{h}_n and ξ_n (as defined above) are satisfied. The claim then follows by Theorem C.1 in [7]. First, we have already observed that $\xi_n(\Omega_n) \subseteq \mathbf{E}_n$ and $\xi_0(\Omega_0) \subseteq \mathbf{E}_0$. Second, we have seen in the proof of part (i) that $h_n(\xi_n)$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable, $n \in \mathbb{N}$. By assumption (f) the extended map $\dot{H}_{\mathcal{S}} : \mathbf{E} \rightarrow \widetilde{\mathbf{E}}$ is $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable, which implies that $\dot{H}_{\mathcal{S}}(\xi_n)$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable. Thus, $\bar{h}_n(\xi_n) = (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$ -measurable (to see this note that, in view of $\widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ = \sigma(\pi_1, \pi_2)$ for the coordinate projections π_1, π_2 on $\widetilde{\mathbf{E}} = \widetilde{\mathbf{E}} \times \widetilde{\mathbf{E}}$, Theorem 7.4 of [2] shows that the map $(h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$ -measurable if and only if the maps $h_n(\xi_n) = \pi_1 \circ (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ and $\dot{H}_{\mathcal{S}}(\xi_n) = \pi_2 \circ (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ are $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable). In particular, the map $\bar{h}_n(\xi_n) = (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$ -measurable, $n \in \mathbb{N}$. Third, we have seen in the proof of part (i) that the map $h_0 = \dot{H}_{\mathcal{S}}$ is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable. Thus the map \bar{h}_0 is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$ -measurable (one can argue as above) and in particular $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ)$ -measurable. Fourth, condition (a) of Theorem C.1 in [7] holds by assumption (b). Fifth, condition (b) of Theorem C.1 in [7] is ensured by assumption (d).

and the continuity of the extended map $\dot{H}_{\mathcal{S}}$ at every point of \mathbf{E}_0 (recall assumption (f)). Hence, (38) holds.

By assumption (g) and the ordinary Continuous Mapping theorem (cf. [8, Theorem 6.4]) applied to (38) and the map $h : \widetilde{\mathbf{E}} \rightarrow \widetilde{\mathbf{E}}$, $(\tilde{x}_1, \tilde{x}_2) \mapsto \tilde{x}_1 - \tilde{x}_2$, we now have

$$h_n(a_n(\widehat{T}_n - \theta_n)) - \dot{H}_{\mathcal{S}}(a_n(\widehat{T}_n - \theta_n)) \rightsquigarrow^{\circ} \dot{H}_{\mathcal{S}}(\xi) - \dot{H}_{\mathcal{S}}(\xi),$$

i.e.

$$a_n(H(\widehat{T}_n) - H(\theta_n)) - \dot{H}_{\mathcal{S}}(a_n(\widehat{T}_n - \theta_n)) \rightsquigarrow^{\circ} 0_{\widetilde{\mathbf{E}}}.$$

By Proposition B.4 in [7] we can conclude (37). \square

The following lemma provides a chain rule for uniformly quasi-Hadamard differentiable maps (a similar chain rule with different \mathcal{S} was found in [30]). To formulate the chain rule let $\widetilde{\mathbf{V}}$ be a further vector space and $\widetilde{\mathbf{E}} \subseteq \widetilde{\mathbf{V}}$ be a subspace equipped with a norm $\|\cdot\|_{\widetilde{\mathbf{E}}}$.

Lemma A.5 *Let $H : \mathbf{V}_H \rightarrow \widetilde{\mathbf{V}}_{\widetilde{H}}$ and $\widetilde{H} : \widetilde{\mathbf{V}}_{\widetilde{H}} \rightarrow \widetilde{\mathbf{V}}$ be maps defined on subsets $\mathbf{V}_H \subseteq \mathbf{V}$ and $\widetilde{\mathbf{V}}_{\widetilde{H}} \subseteq \widetilde{\mathbf{V}}$ such that $H(\mathbf{V}_H) \subseteq \widetilde{\mathbf{V}}_{\widetilde{H}}$. Let \mathbf{E}_0 and $\widetilde{\mathbf{E}}_0$ be subsets of \mathbf{E} and $\widetilde{\mathbf{E}}$ respectively. Let \mathcal{S} and $\widetilde{\mathcal{S}}$ be sets of sequences in \mathbf{V}_H and $\widetilde{\mathbf{V}}_{\widetilde{H}}$ respectively, and assume that the following three assertions hold.*

- (a) *For every $(\theta_n) \in \mathcal{S}$ we have $(H(\theta_n)) \in \widetilde{\mathcal{S}}$.*
- (b) *H is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}} : \mathbf{E}_0 \rightarrow \widetilde{\mathbf{E}}$, and we have $\dot{H}_{\mathcal{S}}(\mathbf{E}_0) \subseteq \widetilde{\mathbf{E}}_0$.*
- (c) *\widetilde{H} is uniformly quasi-Hadamard differentiable w.r.t. $\widetilde{\mathcal{S}}$ tangentially to $\widetilde{\mathbf{E}}_0\langle\widetilde{\mathbf{E}}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{\widetilde{H}}_{\widetilde{\mathcal{S}}} : \widetilde{\mathbf{E}}_0 \rightarrow \widetilde{\mathbf{E}}$.*

Then the map $T := \widetilde{H} \circ H : \mathbf{V}_H \rightarrow \widetilde{\mathbf{V}}$ is uniformly quasi-Hadamard differentiable w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$, and the uniform quasi-Hadamard derivative $\dot{T}_{\mathcal{S}}$ is given by $\dot{T}_{\mathcal{S}} = \dot{\widetilde{H}}_{\widetilde{\mathcal{S}}} \circ \dot{H}_{\mathcal{S}}$.

Proof Obviously, since $H(\mathbf{V}_H) \subseteq \widetilde{\mathbf{V}}_{\widetilde{H}}$ and \widetilde{H} is associated with trace $\widetilde{\mathbf{E}}$, the map $\widetilde{H} \circ H$ can also be associated with trace $\widetilde{\mathbf{E}}$.

Now let $((\theta_n), x, (x_n), (\varepsilon_n))$ be a quadruple with $(\theta_n) \in \mathcal{S}$, $x \in \mathbf{E}_0$, $(x_n) \subseteq \mathbf{E}$ satisfying $\|x_n - x\|_{\mathbf{E}} \rightarrow 0$ as well as $(\theta_n + \varepsilon_n x_n) \subseteq \mathbf{V}_H$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \rightarrow 0$. Then

$$\left\| \dot{\widetilde{H}}_{\widetilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\widetilde{H}(H(\theta_n + \varepsilon_n x_n)) - \widetilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\widetilde{\mathbf{E}}}$$

$$= \left\| \dot{\tilde{H}}_{\tilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\tilde{H}\left(H(\theta_n) + \varepsilon_n \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n}\right) - \tilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\tilde{\mathbf{E}}}.$$

Note that by assumption $H(\theta_n) \in \tilde{\mathbf{V}}_{\tilde{H}}$ and in particular $(H(\theta_n)) \in \tilde{\mathcal{S}}$. By the uniform quasi-Hadamard differentiability of H w.r.t. \mathcal{S} tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\tilde{\mathbf{E}}$

$$\lim_{n \rightarrow \infty} \left\| \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n} - \dot{H}_{\mathcal{S}}(x) \right\|_{\tilde{\mathbf{E}}} = 0.$$

Moreover $(H(\theta_n + \varepsilon_n x_n) - H(\theta_n))/\varepsilon_n \in \tilde{\mathbf{E}}$ and $\dot{H}_{\mathcal{S}}(x) \in \tilde{\mathbf{E}}_0$, because H is associated with trace $\tilde{\mathbf{E}}$ and $\dot{H}_{\mathcal{S}}(\mathbf{E}_0) \subseteq \tilde{\mathbf{E}}_0$. Hence, by the uniform quasi-Hadamard differentiability of \tilde{H} w.r.t. $\tilde{\mathcal{S}}$ tangentially to $\tilde{\mathbf{E}}_0\langle\tilde{\mathbf{E}}\rangle$ we obtain

$$\lim_{n \rightarrow \infty} \left\| \dot{\tilde{H}}_{\tilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\tilde{H}\left(H(\theta_n) + \varepsilon_n \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n}\right) - \tilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\tilde{\mathbf{E}}} = 0.$$

This completes the proof. \square

References

- [1] Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability*, 7, 47–71.
- [2] Bauer, H. (2001). *Measure and integration theory*. de Gruyter, Berlin.
- [3] Belloni, A., Chernozhukov, V., Frenández-Val, I. and Hansen, C. (2016). Program evaluation and causal inference with high-dimensional data, *Econometrica*, to appear (arXiv:1311.2645v6).
- [4] Beutner, E., Wu, W.B. and Zähle, H. (2012). Asymptotics for statistical functionals of long-memory sequences. *Stochastic Processes and their Applications*, 122, 910–929.
- [5] Beutner, E. and Zähle, H. (2010). A modified functional delta method and its application to the estimation of risk functionals. *Journal of Multivariate Analysis*, 101, 2452–2463.
- [6] Beutner, E. and Zähle, H. (2012). Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18, 803–822.
- [7] Beutner, E. and Zähle, H. (2016). Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electronic Journal of Statistics*, 10, 1181–1222.
- [8] Billingsley, P. (1999). *Convergence of probability measures*. Wiley, New York.

- [9] Bühlmann, P. (1994). Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics*, 22, 995–1012.
- [10] Bühlmann, P. (1995). The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Processes and their Applications*, 58, 247–265.
- [11] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap methods and their application*. Cambridge University Press, Cambridge.
- [12] Dudley, R.M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois Journal of Mathematics*, 10, 109–126.
- [13] Dudley, R.M. (1967). Measures on non-separable metric spaces. *Illinois Journal of Mathematics*, 11, 449–453.
- [14] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1–26.
- [15] Efron, B. and Tibshirani, R. (1994). *An introduction to the bootstrap*. Chapman & Hall, New York.
- [16] Gribkova, N. (2002). Bootstrap approximation of distributions of the L-statistics. *Journal of Mathematical Sciences*, 109, 2088–2102.
- [17] Gribkova, N. (2016). Personal communication.
- [18] Krättschmer, V., Schied, A. and Zähle, H. (2015). Quasi-Hadamard differentiability of general risk functionals and its application. *Statistics and Risk Modeling*, 32, 25–47.
- [19] Krättschmer, V. and Zähle, H. (2016). Statistical inference for expectile-based risk measures. *Scandinavian Journal of Statistics*, to appear (arXiv:1601.05261).
- [20] Lahiri, S.N. (2003). *Resampling methods for dependent data*. Springer-Verlag, New York.
- [21] Naik-Nimbalkar, U.V. and Rajarshi, M.B. (1994). Validity of blockwise bootstrap for empirical processes with stationary observations. *Annals of Statistics*, 22, 980–994.
- [22] Peligrad, M. (1998). On the blockwise bootstrap for empirical processes for stationary sequences. *Annals of Probability*, 26, 877–901.
- [23] Pitts, S.M. (1994). Nonparametric estimation of compound distributions with applications in insurance. *Annals of the Institute of Mathematical Statistics*, 46, 537–555.

- [24] Pollard, D. (1984). *Convergence of stochastic processes*. Springer-Verlag, New York.
- [25] Rio, E. (1995). A maximal inequality and dependent Marcinkiewicz–Zygmund strong laws. *Annals of Probability*, 23, 918–937.
- [26] Rubin, D. (1981). The Bayesian bootstrap. *Annals of Statistics*, 9, 130–134.
- [27] Shao, J. and Tu, D. (1995). *The jackknife and bootstrap*. Springer-Verlag, New York.
- [28] Shorack, G.R. and Wellner, J.A. (1986). *Empirical processes with applications to statistics*. Wiley, New York.
- [29] Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak convergence and empirical processes*. Springer-Verlag, New York.
- [30] Varron, D. (2015). Personal communication.
- [31] Zähle, H. (2014). Marcinkiewicz–Zygmund and ordinary strong laws for empirical distribution functions and plug-in estimators. *Statistics*, 48, 951–964.